THE INFLUENCE OF THERAPIST TRAINING AND INTERPERSONAL SKILLS ON CLIENT EMOTIONAL EXPERIENCING

A thesis presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Master of Science

David T. Weibel

November 2003
This thesis entitled

THE INFLUENCE OF THERAPIST TRAINING AND INTERPERSONAL SKILLS ON CLIENT EMOTIONAL EXPERIENCING

BY

DAVID T. WEIBEL

has been approved for
the Department of Psychology and
the College of Arts and Sciences by

Timothy M. Anderson
Associate Professor of Psychology

Leslie A. Flemming
Dean, College of Arts and Science
This study uses the archives from the Ohio University Helping Relationship Study (OUHRS) with hopes of understanding how therapist Facilitative Interpersonal Skills (FIS) influence client emotional experiencing (Klein, Mathieu, Gendlin, & Kiesler, 1969). The OUHRS sought to differentiate the effects of specific or technical skills, which are often associated with training programs, with more general interpersonal skills. The OUHRS study compared psychotherapy outcomes of trained (students with 2+ years in the Ohio University clinical psychology Ph.D. program) and untrained (graduate students from other disciplines) therapists, who also differed in their level of Facilitative Interpersonal Skills (FIS). The FIS measure encompasses the primary interpersonal skills that should contribute to outcome regardless of the particular orientation of the therapist. The results of the OUHRS showed that therapist level of training had no effect on outcome while therapist FIS did.

The current study predicts that high FIS therapists will facilitate higher client emotional experiencing. Eight-minute segments from sessions 2, 3, 4, and 6 were coded using the Experiencing Scale for every client in the study. Some of the results were not as predicted. Therapist FIS did not influence client experiencing. Therapist level of training, which did not significantly influence outcome in the original study, did significantly
influence client experiencing. Possible implications of these findings for the therapeutic process and therapist training are discussed.

Approved: Timothy Anderson

Associate Professor of Psychology
Acknowledgments

I would like to thank Timothy Anderson for his guidance, support, and generosity. I am grateful to everyone who contributed to the OUHRS study, providing me with such a rich source of data. I thank Ben Ogles and John Garske for their contributions and advice. Lastly I thank my family for standing by me as I searched for a career that would provide me with a greater sense of meaning.
# Table of Contents

Abstract ................................................................................................................................. 3  
Acknowledgments ............................................................................................................. 5  
List of Tables ....................................................................................................................... 7  
List of Figures ...................................................................................................................... 8  
Introduction ......................................................................................................................... 9  
  The History of Psychotherapy Research ........................................................................ 16  
  Common Factors Theories ............................................................................................ 24  
  Therapist Effects ........................................................................................................... 45  
  The OUHRS ................................................................................................................... 51  
  The Current Study: Linking Skill to Process ............................................................... 53  
  Hypotheses .................................................................................................................... 54  
Method .................................................................................................................................. 55  
  Sample ............................................................................................................................ 55  
  Procedure ....................................................................................................................... 60  
  Measures ........................................................................................................................ 71  
Results .................................................................................................................................. 81  
Discussion ........................................................................................................................... 88  
  Limitations ...................................................................................................................... 115  
  Conclusion ...................................................................................................................... 119  
References .......................................................................................................................... 121  
Appendix A: Coder Debriefing Responses ....................................................................... 138
List of Tables

Table

1. OUHRS design...................................................................................................................... 52
2. Correlations between each coder and Klein manual weeks 1-3 ........................... 67
3. Correlations between each coder and Klein manual final assignments.............. 67
4. Correlations between coders for mode and peak scores ........................................ 69
5. Correlations between each coder and all other coders........................................... 69
6. Correlations between coders after removal of coder L........................................ 70
7. Intercorrelations between experiencing and hypothesized variables............... 82
8. Experiencing scores for FIS, training status and client social skill .................. 85
9. MANOVA for modes ......................................................................................................... 86
10. MANOVA for peaks.......................................................................................................... 87
11. Correlations between modes of different segments .............................................. 116
12. Correlations between peaks of different segments .............................................. 117
13. Mean Experiencing scores for trained and untrained therapists ..................... 118
List of Figures

Figure

1. Frequency of mode and peak ratings given by coders........................................ 90
The Influence of Therapist Training and Interpersonal Skills
On Client Emotional Experiencing

INTRODUCTION

This study used the archives from the Ohio University Helping Relationship Study (OUHRS) with hopes of understanding how therapist Facilitative Interpersonal Skills (FIS) influence client emotional experiencing (Klein, Mathieu, Gendlin, & Kiesler, 1969). The OUHRS sought to differentiate the effects of specific or technical skills, which are often associated with training programs, with more general interpersonal skills. The OUHRS study compared psychotherapy outcomes of trained (students with 2+ years in the Ohio University clinical psychology Ph.D. program) and untrained (graduate students from other disciplines) therapists who also differed in their level of Facilitative Interpersonal Skills (FIS). FIS (Anderson, 2001a) is based on the psychotherapy literature and encompasses the primary skills that should contribute to outcome regardless of the particular orientation of the therapist. The FIS rating was based on the Social Skills Inventory (Riggio, 1989) and the Facilitative Interpersonal Skills Test (see Measures section).

First I will delineate what is meant by specific and common factors in this paper. There are varied interpretations of these terms befitting the complex and subjective nature of psychotherapy. According to Jones, Cumming, and Horowitz (1988) specific factors are “well-defined, intentional actions on the part of the therapist, such as interpretation, labeling of feeling, or the correction of distorted beliefs” (p. 48). Specific factors are
allied with a specific therapy or orientation (Kivlighan & Schmitz, 1992) and would likely be discussed in a therapy manual.

Specific factors are usually operationalized as response modes (Elliot et al., 1987) or therapist intentions (Hill & O'Grady, 1985). Psychotherapists have traditionally asserted that the technical operations stressed by their particular “school” or “orientation” are the critical ingredients responsible for change (Strupp & Hadley, 1979). Evidence is presented later that this may not be the case.

Non-specific factors are those elements that are not techniques, may or may not contribute to therapy and are often not covered in the treatment manual. Non-specific factors include everything that is not a technique. This would include such things as visiting an office, sitting in chairs and paying a fee. Common factors are those elements that are common to all (or most) therapies and that are responsible for psychotherapeutic benefits (Hubble, Duncan, & Miller, 1999). Examples would be the therapeutic alliance, empathy, and emotional experiencing. There may not always be a clear distinction between what is common and what is specific. Undoubtedly there are interactions between the two (Arkowitz, 1997). A therapist delivers specific techniques within the context of the alliance, a common factor, using qualities such as warmth and understanding, also common factors. It would also be difficult to build the alliance or express empathy without using specific techniques.

Many of these factors can be influenced by the therapist, and some models of common factors include therapist contributions (Frank & Frank, 1991; Grencavage & Norcross, 1990). The word common does not mean that all therapists have these abilities in equal
amounts. Studies have shown that therapists vary in their ability to form alliances (Luborsky, McLellan, Woody, O'Brien, & Auerbach, 1985), show interest in the client (Luborsky, McLellan, Diguer, Woody, & Seligman, 1997), or show warmth and understanding towards the client (Najavits & Strupp, 1994). Common simply means that these factors facilitate many different types of therapy.

Common factors may also be more important in particular therapies or at particular points over the course of therapy. Common factors may contribute to success across different therapies, but that does not mean their use and importance does not vary depending on the therapy or the moment within that therapy (Arkowitz, 1997). Some approaches such as existential may go so far as to say the ‘relationship is the therapy’ (Yalom, 2002), while behavioral approaches have not always focused on the relationship and more recently have seen the relationship as facilitating techniques (Follette, Naugle, & Callaghan, 1996). Arkowitz (1997) argues that common factors such as the alliance should be treated as processes which can play different roles across different therapies or over time in the same therapy. For example, early in all therapies it is necessary to establish rapport. In later stages the relationship may be the source of change in some instances (e.g., corrective emotional experience), serve as a source of identification for the client, or serve as a motivator for change in others. The alliance may do all of these at different times within the same therapy.

Research interest in common factors gained momentum when comparisons between treatments revealed similar outcomes (Elkin, 1994; Lipsey & Wilson, 1993; Luborsky & Singer, 1975; Shapiro et al., 1994; Smith, Glass, & Miller, 1980; Stiles, Shapiro, & Elliot,
This caused many researchers to search for possible common elements between the treatments separate from the level of techniques that could be contributing to outcomes. The OUHRS was an attempt to examine the contributions of specific versus common factors. The FIS construct was designed to encapsulate a therapist’s ability to provide those factors which have most often been described as common, such as a strong alliance, empathy and warmth. Including trained therapists would hopefully demonstrate if specific techniques learned in graduate school contribute to outcome. Another possibility is that graduate school would improve the ability to deliver common factors such as empathy or build alliances, though this does not seem to have been a major focus of most training programs to date (Alberts & Edelstein, 1990; Ford, 1979; Stein & Lambert, 1995).

The results of the OUHRS indicated that the level of training did not effect outcome, but therapist FIS did. High FIS therapists also had higher client-rated therapeutic alliances than low FIS therapists. While we know that the high FIS therapists facilitated better alliances and higher outcomes, we do not know how they influenced the actual sessions. This leads to the current study which aimed to show that one of the processes that FIS may have influenced is client emotional experiencing. Specifically, this study predicted that high FIS therapists would facilitate higher levels of client experiencing.

In order to understand the relevance of the OUHRS and this thesis, some of the history of psychotherapy research is presented. Throughout the history of psychotherapy research, there has been a relative lack of emphasis on therapist skill. Therapist personal
characteristics have not been the subject of intense investigation due to numerous forces within the field. Psychotherapy researchers first needed to establish the efficacy of psychotherapy in the face of challenges by those such as Eysenck (1952). Meeting this challenge entailed focusing on therapies as a whole and not the contributions of individual therapists. Once this challenge was met the majority of research efforts went towards determining which therapy was better. Currently, economic pressures are forcing psychotherapy researchers to provide empirical support for treatments so that health maintenance organizations deem them reimbursable. The dominant research model that has emerged is the randomized clinical trial (RCT), which necessitates that therapist differences be minimized and treated as error variance.

A very broad overview of RCT’s and how they take the emphasis away from therapist and common factors is offered. Interestingly, the huge investment in RCT’s has failed to differentiate between technically diverse therapies, with a few exceptions. Ironically, therapist differences, which RCT’s attempt to minimize, often account for more outcome variance than the treatments.

The situation of equal outcomes given diverse therapies has been commonly referred to as the Dodo bird verdict. One way to explain the Dodo bird verdict is the common factors theory, which focuses on curative elements common to all therapies. Several theories or models have been proposed to explain common factors. These will be reviewed and similarities between them will be described. Particular focus will be placed on factors that can be influenced by the therapist. Three common factors or processes that are present in most models of common factors are the relationship or therapeutic alliance,
therapist qualities such as warmth and empathy, and the chance to express emotions or look inwards.

Studies have provided evidence that the therapist can influence these three common factors and that they are somewhat intertwined. Therapist qualities such as warmth and empathy have been shown to contribute to the working alliance. Both therapist qualities and the working alliance have been shown to effect the level of emotional experiencing of the client. Thus, the process of emotional experiencing is a good candidate for a process that FIS would influence. These therapist-influenced common factors may also help us understand the therapist effects which appear in RCT’s.

After explaining how common factors can help explain the Dodo bird puzzle, we will return to the therapist effects that were found in the RCT’s. Evidence from major RCT’s will be presented which reveals that therapist effects tend to be larger than differences attributable to therapies and techniques. Then studies which attempt to elucidate therapist effects by measuring what more and less successful therapists do differently will be explored. Interestingly enough, when these studies are able to differentiate what therapists do differently, it is usually the therapists’ general skills that distinguish the therapists’ performance. A number of studies will be presented which provide evidence that the therapist’s ability to utilize and provide common factors can distinguish the therapists’ performance. The present study is based on the assumption that the therapist’s ability to provide and utilize common factors, which probably depends on their interpersonal skills, is a significant therapeutic factor.
The OUHRS study revealed that FIS did influence the building of alliances and therapeutic outcome. However, it is still not clear exactly how FIS influenced moment-to-moment processing within sessions. One such measure of therapeutic process is the emotional experiencing scale. Experiencing refers to the quality of a person’s self-involvement and participation in therapy. It refers specifically to the extent to which inner referents become the felt data of attention, and the degree to which efforts are made to focus on, expand, and probe those data. Experiencing refers to the quality of an individual’s experiencing of himself or herself, the extent to which his or her ongoing, bodily, felt flow of experiencing is the basic datum of his or her awareness and communications about himself or herself, and the extent to which this inner datum is integral to action and thought. The scale is not dependent on the therapist’s orientation or technique. Experiencing should be relevant for all therapies where change in the patient’s level of expressiveness, self-awareness, or self understanding is a goal, or where self-attitudes and expressive behavior are in any way manipulated as part of the treatment procedure (Klein et al., 1969).

The common factors section of this paper will reveal that experiencing or related constructs such as emotional involvement are often listed as a common factor, important for therapeutic success in a variety of orientations or therapies. Evidence will be presented that experiencing is influenced by the skills that went into the construction of the FIS scale, such as warmth and empathy. Experiencing has also been found to be higher when the alliance is higher, and the high FIS therapists did have stronger alliances
in the OUHRS. The main hypothesis of the current study is that higher therapist FIS will lead to higher experiencing on the part of the client.

The History of Psychotherapy Research: Moving Towards the Current Study

Comparing treatments, not therapists

Psychotherapy researchers have not placed a huge emphasis on studying therapist skills in the past. Moncher and Prinz (1991) reviewed 359 treatment outcome studies and concluded that most ignored or failed to adequately assess therapist competence or skill as a variable in treatment efficacy. Part of this stems from the fact that psychotherapy researchers have needed to meet several challenges that precluded them from focusing on skills. First researchers felt the need to respond to the gauntlet that had been thrown down by Eysenck (1952). Eysenck’s literature review stated that the benefits provided by psychotherapy were not better than the spontaneous remission rate for neuroses. Over time psychotherapy researchers succeeded in demonstrating the efficacy of psychotherapy (Kendall & Chambliss, 1998; Lambert & Bergin, 1994; Lipsey & Wilson, 1993; Seligman, 1995). This was an important challenge for the field to meet, because it ensured that psychotherapy would be reimbursable and encouraged more research to be done.

Then much of psychotherapy researchers’ efforts focused on determining which orientation or therapy was superior. This movement has gained strength recently as health maintenance organizations have indicated that they will only reimburse therapies that have been empirically supported. The research paradigm which emerged to compare
therapies and gain support for them is the random clinical trial (RCT), which was borrowed from the world of medicine.

RCT’s attempt to control for any variable which can interfere with the comparison of two treatments. Researchers attempt to standardize client variability by using homogeneous samples. The clients are screened to meet criteria for specific disorders as delineated in the latest version of the DSM, and they are excluded if they have comorbid conditions. RCT’s attempt to standardize therapists by selecting therapists with an equal amount of training and then instructing them to use manuals to provide a pure “dose” of the therapy. The aim is to eliminate the therapist as a variable that might account for client improvement. Researchers often treat any variation in therapist performance as error. These efforts to standardize the clients and therapists are done so that that the specific effects of the treatments can be identified (Talley, Strupp, & Butler, 1994).

RCT’s have been successful in meeting the challenge posed by Eysenck and establishing the efficacy of various therapies. RCT’s can be excused for not studying therapist effects in detail, because that was not their purpose and major funding agencies and the field in general were not calling for such studies. RCT’s are expensive, and it is difficult to recruit enough therapists to compare their performance with sufficient statistical power (Crits-Christoph & Mintz, 1991). The medical professions, from whom psychologists borrowed the RCT model, have never found it necessary to empirically compare the interpersonal skills of various physicians. Thus, it makes sense that psychologists were not racing to compare their performance in therapy.
While it is understandable why RCT’s have attempted to minimize therapist effects, doing so makes them vulnerable to the “Uniformity Myth”, a warning given by Kiesler (1966) against treating all therapists or clients as equivalent. Although RCT’s attempt to eliminate or minimize therapist effects, these effects still appear and often account for more variance than the treatment effects that are the independent variables of these studies (Beutler, Machado, & Neufeldt, 1994; Crits-Christoph, Baranackie, Kurcias, Beck, & al, 1991; Lafferty, Beutler, & Crago, 1989; Orlinsky, Grawe, & Parks, 1994). It is the premise of this study that researching the differences between therapists may be more interesting and beneficial to the field than researching the differences between therapies. In fact many of the RCT’s failed to show major differences between therapies. This contentious issue is discussed in the next session.

The Dodo Bird Verdict: Failure of Specific Techniques to Influence Outcome

Despite repeated attempts to determine which therapies are superior, it appears that outcomes are generally similar among therapies (Elkin, 1994; Lipsey & Wilson, 1993; Luborsky & Singer, 1975; Shapiro et al., 1994; Smith, Glass, & Miller, 1980; Stiles, Shapiro, & Elliot, 1986; Strupp & Hadley, 1979; Wampold et al., 1997; Wampold, 2001). This potentially frustrating situation of equivalent outcomes given varying treatments has been labeled the Dodo Bird verdict, borrowing from Alice’s Adventures in Wonderland. The Dodo stated, “Everybody has won and all must have prizes” (Rosenzweig, 1936). Under this scenario all therapies win and each therapy’s devotees can continue to claim superiority.
Some researchers have taken issue with a blanket acceptance of the Dodo verdict. In response to Wampold’s (1997) article which stated “Empirically, All must have Prizes” in its title, Chrits-Christoph (1997) questioned the type of studies that were included in Wampold’s analysis. He stated that 79 of the 114 studies involved comparisons of cognitive and behavioral treatments for anxiety problems and that similarities between these treatments does not imply similarities between all treatments. He also pointed out that among the 29 studies that were not comparisons of simple, individual cognitive-behavioral treatments with each other, 14 of the 29 reported some meaningful differences between treatment conditions. Chrits-Christoph concludes that “the Dodo verdict may indeed be correct regarding some patient problems, but much more research is needed before it is known whether this verdict applies to the major clinical disorders.” In fact, there is an entire movement which aims to show that for certain disorders there are empirically supported treatments which practitioners should use.

EST’s: have they made the Dodo extinct?

The Empirically Supported Treatment (EST) movement seems to fly in the face of the Dodo Bird verdict. The EST movement, which arose out of the APA division 12 Task Force on the Promotion and Dissemination of Psychological Procedures (Procedures, 1995), was designed to identify treatments with proven efficacy for particular disorders. The task force has identified a number of effective, or probably effective, psychological treatments for a variety of mental disorders including depression, eating disorders, marital discord, panic disorder with and without agoraphobia, specific phobias,
posttraumatic stress disorder, schizophrenia (in combination with medication), smoking cessation and borderline personality disorder (Barlow, Levitt, & Bufka, 1999).

Does this impressive list of treatments invalidate the Dodo bird verdict? In order to understand how the field can have both the EST’s and the Dodo Bird verdict, it is necessary to understand that the researchers who support either position have different criteria and may be looking at different data. Researchers expounding the Dodo bird verdict are summarizing meta-analyses of dozens of studies of head to head comparisons between therapies (Wampold, 2001).

To become an EST requires a different set of criteria. To be deemed a well-established treatment, a treatment must have two group design studies showing that it is superior to a pill or psychological placebo or equal to an established treatment. A large series of single case design experiments (n > 9) demonstrating efficacy can also provide well-established status. The criteria to achieve the status of a probably efficacious treatment are slightly less rigorous (Chambless & Hollon, 1998; Elliot, 1998).

Some authors feel that the EST list shows a pattern of unintended bias against certain classes of treatments such as psychodynamic and experiential treatments which are not as easily manualizable as cognitive and behavioral therapies, do not share the same assumptions regarding the nature of treatment, or place the same emphasis on research (Bohart, O'Hara, & Leitner, 1998; Elliot, 1998; Fonagy & Target, 1996; Henry, 1998). It may be that the cognitive and behavioral treatments are superior, but it may also be that they are more suited to earning the status of an EST. Because a therapy achieves EST status does not necessarily mean that it is superior. In an update to the release of the EST
report, Chambless et. al (1996) stated, “the list is far from complete and should not be employed as the basis for decisions concerning reimbursable treatments by third party payers…That a treatment is not on our list in no way means that it has been shown to be ineffective.” Although the search for EST’s has compiled impressive evidence for a number of efficacious treatments, it has not driven the Dodo verdict to extinction.

The vast majority of head to head comparisons between therapies have yielded little if any differences (Elkin, 1994; Lipsey & Wilson, 1993; Luborsky & Singer, 1975; Shapiro et al., 1994; Smith, Glass, & Miller, 1980; Stiles, Shapiro, & Elliot, 1986; Strupp & Hadley, 1979; Wampold et al., 1997; Wampold, 2001). A further point in favor of the Dodo verdict is that allegiance effects of the researchers conducting these comparisons often appear. One study that evaluated the effects of prominent psychotherapy researchers’ allegiances found that 69% of effect sizes are attributable to the researchers allegiances (Luborsky et al., 1999). Even with these huge allegiance effects, the researchers rarely find significant differences between therapies. In their meta-analysis of outcome studies, Wampold et al. (1997) estimated that the effect sizes of head-to-head comparisons hovered near zero with an approximate range of from 0 to .21. In a later meta-analysis, Wampold and Serlin (2000) examined the effect of improperly modeling nested designs in psychotherapy studies. They believe that if therapist effects are properly modeled, the effect size of .20 for treatments is completely artifactual. They conclude that in reality treatment effects appear to be zero. The belief in the Dodo Bird verdict has led some researchers to focus on why different treatments produce comparable outcomes, rather than continuing to compare therapies without attention to common factors.
Looking to Explain the Dodo: Are techniques truly different?

Some researchers accept the Dodo Bird verdict and have sought to explain why it occurs. One way to explain the Dodo bird verdict would be if the treatments were actually not different. At one point it was believed that the behavior of therapists in session might have been the same despite their theoretical differences (London, 1964). This question has now been answered as process researchers have repeatedly demonstrated that therapists act differently across different therapies, at least at the technical level (DeRubeis et al., 1990; Elliot et al., 1987; Hill, O'Grady, & Elkin, 1992; Hill, Thames, & Rardin, 1979; Stiles, 1979; Stiles, Shapiro, & Firth-Cozens, 1988; Strupp, 1955, 1957). Stiles (1979) used his Verbal Response Mode system to show that therapists of differing orientations use different patterns of response modes. Stiles found that prominent of practitioners from three major therapy schools differed significantly in their use of response modes. Client-centered therapists used the client’s frame of reference (mostly Acknowledgement and Reflection), Gestalt therapists used their own frame of reference (Disclosure, Advisement, Question, and Interpretation) and psychoanalytic therapists used the client’s experience (Question, Acknowledgement, Interpretation, and Reflection) for over 90% of their utterances. Therapists who adhere to different orientations produce different patterns of interventions, and these patterns of interventions have not often led to different outcomes.

This pattern was likewise found in a qualitative analysis of the OUHRS (Anderson, Weibel, Wang, Fende, & Baldrachi, 2001). Coders took detailed process notes of every session for 7 dyads to gain a better understanding of the findings of the study as well as
generate new hypotheses. Five of the dyads were selected because they represented the findings of the study. Thus, three high FIS therapists who achieved good outcomes, as well as two low FIS therapists with poor outcomes, were studied. At a later date two cases whose outcomes did not agree with the main findings were also examined (i.e., high FIS with poor outcome). While the qualitative analyses was designed to understand how FIS was demonstrated in session, the coders also noticed different intervention patterns among the trained and untrained therapists. The untrained therapists offered more self-disclosure and advisement, much as a friend would. The trained therapists offered more restatements, clarifications and interpretations. The original results of the OUHRS revealed that these different intervention patterns between trained and untrained therapists did not influence outcomes. Thus, there is some evidence in the literature and the OUHRS that orientation and training effect therapeutic operations at a technical level without effecting outcome. When specific interventions are looked at by themselves, there also does not seem to be a clear connection to outcome.

The Contribution of Specific Techniques Remains Unclear

Interventions such as giving advice, directiveness, reflection, open-ended questioning, and support (i.e., encouragement) have generally shown a mixed association with outcome (Beutler et. al, 1994). In addition studies of verbal response modes have shown that they account for only about 1% of outcome variance (Elliot, Barker et al. 1982; Elliot 1985; Hill, Helms et al. 1988).
Different types of clients may react differently to intervention patterns, or an intervention’s impact may be mediated by various client or therapist factors, such as interpersonal skill. For example, therapist directiveness was found helpful with resistance-prone depressed clients, whereas nondirective therapists were more successful with low-resistance-prone depressives (Beutler et. al, 1991). In another study exploratory actions were found to better suit clients who were highly motivated or had more coherent self-concepts, whereas supportive interventions were better suited to clients who were less motivated and whose self-concepts were less stable (Horowitz et. al, 1984). The characteristics of the therapist and client, as well as the relationship between them, may effect the psychotherapeutic process in ways that are more significant than the application of specific techniques.

Common Factors Theories

The Dodo Bird verdict, as well as findings that specific techniques do not substantially contribute to outcome, have caused some researchers to look towards common factors that contribute to therapy. The common factors theory posits that it is the elements that are shared between therapies such as therapist warmth, therapist empathy, the placebo effect, the therapeutic alliance, and client factors that are responsible for change.

This is an unwieldy research area as there are a number of different conceptualizations that include different factors. First some theories of common factors will be explored. Then I will point out similarities between the models and focus on those factors that the therapist can influence. The empirical evidence linking these common factors to outcome
will be reviewed, and the ways in which these common factors interact will also be explored.

Models of Common Factors

Frank and Frank (1991) describe six common elements to the rituals and procedures used by all psychotherapists:

1) The therapist combats the client’s sense of alienation by developing a relationship that is maintained after the client divulges feelings of demoralization.

2) The therapist maintains the client’s expectation of being helped by linking hope for improvement to the process of therapy.

3) The therapist provides new learning experiences.

4) The client’s emotions are aroused as a result of therapy.

5) The therapist enhances the client’s sense of mastery or self-efficacy.

6) The therapist provides opportunities for practice.

Castonguay (1993) has elaborated three meanings that can be applied to common factors:

1) Global aspects of therapy that are not specific to any one approach: insight, corrective experiences, opportunity to express emotions, acquisition of a sense of mastery

2) Therapeutic context and therapeutic relationship (the working alliance)

3) Client expectancies and involvement in the therapeutic process

Hans Strupp (1986) defined psychotherapy as the systematic use of a human relationship for therapeutic purposes. The therapist's interventions are guided by a theory
that addresses the nature of the patient's difficulties and factors instrumental in their perpetuation. The therapist’s role is to create an interpersonal context, understand the meaning of the patient's feelings and wishes, and assist with the identification and reformulation of the patient's meanings.

In an attempt to organize the common factors literature, Grencavage & Norcross (1990) reviewed papers which discussed common factors and devised five areas of commonality:

1) **client characteristics** - positive expectation, hope or faith; distressed or incongruent client; patient actively seeks help
2) **therapist qualities** - general positive descriptors; cultivates hope, enhances expectations; warmth, positive regard
3) **change processes** - opportunity for catharsis-ventilation, acquisition and practice of new behaviors, provision of rationale
4) **treatment structures** - use of techniques-rituals, focus on "inner world"-exploration of emotional issues, adherence to theory
5) **relationship elements** - development of alliance-relationship (general); engagement; transference

Lambert & Ogles (in press) break down common factors into 3 categories:

1. Support factors – therapeutic alliance, trust, therapist/client active participation, therapist warmth/respect/empathy/acceptance/genuineness, catharsis
2. Learning factors – advice, insight, corrective emotional experience, affective experiencing, cognitive learning

Lambert and Ogles’ categories represent the presumed developmental sequence that operates in psychotherapy (i.e. support precedes changes in beliefs/attitudes, which precede attempts of therapist to encourage patient action). Together, the common factors provide a cooperative environment where the patient’s increased sense of trust and safety, along with decreases in tension and anxiety lead to changes in conceptualizing problems and acting via reframing fears, taking risks, and working through interpersonal relationship problems.

These various common factors theories make it clear that there is not a consensus about what constitutes a common factor. On the other hand, there are parallels between these conceptualizations of common factors. This thesis will focus on the common factors which were listed by all of these authors, have good empirical support, and relate to the contribution of the therapist. Three of the commonalities between these theories are the relationship or working alliance, therapist provided factors (warmth and empathy), and the process of emotional expression or looking inwards. These three common factors have empirical evidence behind them. They also seem to be linked. Therapist empathy and the alliance are often found to be correlated. Therapist provided factors such as empathy and the alliance have been found to increase emotional experiencing. The empirical evidence behind these three common factors will now be reviewed.
Empirical Evidence Behind Common Factors

The therapist attitudes identified by Rogers and colleagues (Rogers, Gendlin, Kiesler, & Truax, 1967; Truax & al, 1966) in the late 1950’s and 1960’s remain important ingredients of a positive therapy relationship, especially from the vantage point of clients. Roger’s idea that accurate empathy, non-possessive warmth, and genuineness represent the “necessary and sufficient” conditions of beneficial outcome has generated an extensive body of studies that dominated research on the relationship for more than three decades (Bachelor & Horvath, 1999). In a review of the conclusions reached by nine major review articles on relationship variables, Patterson (1984) concluded that “there are few things in the field of psychology for which the evidence is so strong” as that supporting the “necessity if not sufficiency, of the therapist conditions of accurate empathy, respect, or warmth, and therapeutic genuineness.” p (437). Most schools now accept that these are fundamental in the formation of the working alliance and necessary for progress in psychotherapy. These findings are repeated in a number of reviews of the literature (Beutler et al., 1994; Lambert & Bergin, 1983; Orlinsky & Howard, 1986)

Empathy

A study by Lafferty et. al (1989) examined differences between more and less effective trainee psychotherapists. Therapists were assigned to one of two groups depending on whether the preponderance of their patients' changes in symptomatology indicated more or less improvement over the course of therapy. The authors used a stepwise discriminant function procedure to see if therapist variables such as emotional
adjustment, relationship skills, eliciting patient involvement, credibility, directiveness, and theoretical orientation would differentiate between the effective and less effective therapists. Empathy was the variable which was most predictive of being an effective or ineffective therapist. Patient involvement was the second variable to enter the function, followed by therapist directiveness/support. The discriminant function equation accurately classified 78.57% of all therapists. Less effective therapists were revealed to have lower levels of empathic understanding, to rate their patients as more involved in treatment, and to rate themselves as more supportive than the more effective therapists.

Empathy has even been cited as important in the therapies of orientations that traditionally have not focused on empathy (Burns & Nolen-Hoeksema, 1992). Examining the effects of empathy on a group of 185 depressed individuals (aged 18-75 yrs) treated with cognitive-behavioral therapy, Burns and Nolen-Hoeksema found that the patients of the therapists who were the warmest and most empathetic improved significantly more than the patients of the therapists with the lowest empathy ratings, when controlling for initial depression severity, homework compliance and other factors. Using structural equation modeling, this study demonstrated that therapeutic empathy has a moderate-to-large causal effect on recovery. Therapeutic empathy predicted final BDI scores in all five models the authors ran (p <.01 in all five). The authors simultaneously estimated the reciprocal effects of depression severity on therapeutic empathy and found that this effect was quite small. Thus, therapists were not being more empathic because their clients were becoming more depressed. Subjects who terminated therapy prematurely were less likely to complete the self-help assignments between sessions, rated their therapists as
significantly less empathic, and improved significantly less. Subjects with borderline personality disorder improved significantly less, but they rated their therapists as just as empathic and caring as other patients.

Empathy can also show important contributions to outcome even in studies not specifically designed to show its importance. Miller, Taylor and West (1980) investigated the comparative effectiveness of various behavioral approaches aimed at helping problem drinkers control their alcohol consumption. Although the focus of the study was on comparing the different therapies, they collected data on the contributions of therapist empathy to patient outcome. The authors found a strong relationship between empathy and outcome obtained from 6 to 8 month follow-up interviews used to assess drinking behavior. Therapists’ rank on empathy correlated ($r = .82$) with patient outcome, thus accounting for 67% of the variance in the criteria.

Other Studies on Empathy, Warmth and Positive Regard

A study by Benunn and Schindler (1988) investigated the specific therapist and client factors operative within a cognitive-behavioral treatment of 35 adult phobic patients. Clients received 12 weekly sessions of cognitive-behavioral therapy (CBT) that included exposure and cognitive interventions. Traditionally CBT has been a good example of a treatment that ascribes it success to the effective delivery of specific techniques. The three therapist factors which contributed to outcome were positive regard, competence/experience and activity/guidance. These factors distinguished between participants who benefited from treatment. Even in a highly technique driven therapy,
common factors such as positive regard contribute significantly to the therapeutic process.

Bryant et. al (1999) examined the contributions of therapist skill to homework compliance. The cognitive therapy clients who were more compliant with homework exhibited significantly greater treatment response on one depression measure, but not another. The researchers broke cognitive therapy skills into general therapeutic skills and cognitive therapy-specific components. The general therapeutic skills subscale included items of “Understanding”, “Interpersonal Effectiveness,” and “Collaboration.” The cognitive therapy subscale included the items “Empiricism/Guided Discovery,” “Focusing on Key Cognitions or Behaviors,” “Strategy for Change,” and “Application of Cognitive-Behavioral Techniques.” General therapeutic skills were predictive of homework compliance, but cognitive therapy specific skills were not. Again, general or common skills were more important than specific skills.

Many other studies have shown the importance of common factors such as acceptance and understanding. Lorr (1965) asked 523 psychotherapy clients to describe their therapists on 65 different statements. A factor analysis of these data identified five factors: understanding, accepting, authoritarian (directive), independence encouraging, and critical hostile. Patient ratings of understanding and accepting correlated most highly with patient and therapist rated improvement.

Elliot and James (1989) reviewed 9 domains of client experience that appear in the existing empirical literature. Five of these pertain to clients' experiences of their own psychological processes during therapy: intentions, feelings, style of self-relatedness,
style of relating to therapist, and central concerns. Two domains refer to clients' experiences of the action and person of their therapists; the final two domains concern clients' experiences of change in therapy. The aspects of therapy most frequently reported by clients as helpful were interpersonal themes including allowing client self-expression, a supportive relationship, and facilitative therapist characteristics such as warmth.

Paulson et. al (1999) examined what clients experience as helpful in counseling via concept mapping, a methodological approach combining qualitative and quantitative strategies. The purpose of the study was to clarify the scope and interrelations among elements of the retrospective experience of helpfulness among 36 clients who had completed counseling after a mean of 11 sessions. The first cluster of items which the clients found helpful about counseling concerned what they termed counselor facilitative interpersonal style.

Therapeutic Alliance

The therapeutic alliance is the most discussed and empirically supported common factor (Bachelor & Horvath, 1999; Henry & Strupp, 1994; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000). In recent years, interest in the therapeutic alliance has increased. The therapeutic alliance was first described by Freud (1958, 1966). He underscored both the importance of the analysand’s attachment to the psychoanalyst and the psychoanalyst’s interest in and “asympathetic understanding” of the patient in the early treatment relationship.
Gaston (1990) later attempted to define the components of the alliance. She suggested that the following components were measured by alliance ratings scales: (a) the client’s affective relationship to the therapist, (b) the client’s capacity to work purposefully in therapy, (c) the therapist’s empathic understanding and involvement, and (d) the client-therapist agreement on the goals and tasks of therapy. Therapists aim to use their personal skills to build an alliance and thus improve outcomes. Researchers such as Luborsky (1976) and Bordin (1994) argued that the alliance, viewed as a positive reality-based component of the therapeutic relation, was ubiquitous and universal in all successful helping endeavors. Bordin identified three components of the therapeutic alliance: bonds, tasks, and goals.

There is currently strong agreement that the therapeutic relationship is an important component in all forms of therapy, and that the quality of the relationship influences the outcome of therapy. Research has shown that the quality of the therapeutic relationship or related constructs such as the “therapeutic bond” influence outcome across a variety of treatment environments and across a wide range of client problems (Beutler et al., 1994; Horvath & Symonds, 1991; Luborsky, Crits-Christoph, Mintz, & Auerbach, 1988; Orlinsky & Howard, 1986). However, the use of the relationship and its salient components may differ (Callaghan, Naugle, & Follette, 1996; Gomes-Schwartz, 1978; Raue, Castonguay, & Goldfried, 1993). The therapeutic alliance has also been shown to play an important role across therapies of different orientations such as behavioral, eclectic, and psychodynamic therapies (Gaston, Marmar, Thompson, & Gallagher, 1988; Horvath, 1994). The alliance may be an important factor in group marital therapy
(Bourgeois, Sabourin, & Wright, 1990; Pinsof, 1994) and may contribute to relations not structured as psychotherapy such as pharmacotherapy with minimal supportive contact (Krupnick, Stotsky, Simmens, & Moyer, 1992, June).

The alliance is not only a product of therapeutic success, meaning that clients do not only give high alliance ratings after therapy is going well. There are data that suggest the relationship quality is an active factor, contributing to the success of therapy over and above concurrent therapeutic gains (Gaston, Marmar, Thompson, & Gallagher, 1991). Thus the relationship can produce change, and is not only a reflection of beneficial results (Lambert & Bergin, 1994).

It seems that the quality of the alliance early in therapy (i.e., the third to fifth session) is most predictive of outcome. Although the relation between the alliance measured at later stages is still significant, it is more modest. There is also evidence that the alliance can fluctuate for individual clients over the course of therapy (Horvath & Marx, 1988, April, 1990; Safran, Crocker, McMain, & Murray, 1990; Safran, Muran, & Wallner Samsta, 1994).

Clients and therapists differ in their perceptions of the therapeutic relationship. Client and therapist ratings of the quality of the relationship have consistently shown low agreement (Golden & Robbins, 1990; Gurman, 1977; Horvath & Marx, 1990; Tichenor & Hill, 1989). Therapists must not assume that their evaluation of the quality of the relationship matches their clients’ perceptions. Nor can they assume that the client understands their interventions or intentions to form a positive relationship (Hill et al., 1988; Hill & O'Grady, 1985; Horvath, Marx, & Kaman, 1990). Given these differential
evaluations, it seems essential that therapists have the skills to check with the client and attempt to verify that and clarify disparate understandings of the events that may threaten the therapeutic relationship.

It also appears that therapists and clients value different aspects of the relationship. Therapists tend to emphasize the role of the clients’ contributions (e.g., active participation in the therapy process) in client change, perhaps because therapists view clients as the agents of their own improvement, whereas clients tend to value therapist characteristics such as therapist-provided help and demonstrated warmth, caring and emotional involvement (Bachelor, 1991, 1995; Lambert & Bergin, 1983; Lazarus, 1971; Murphy, Cramer, & Lillie, 1984). The client’s perceptions of the relationship have been shown to be superior to the therapists in predicting outcome (Horvath & Symonds, 1991; Lambert & Bergin, 1994), and clients appear to be good judges of potentially curative factors of the relationship (Bachelor, 1991; Grigg & Goodstein, 1957; Murphy et al., 1984). Thus, therapists must attend to the clients’ perceptions of the relational climate and ensure that factors deemed valuable by clients are actually communicated and delivered to them.

Relationship factors have also been found to be important in group therapy. Glass and Arnkoff (1988) examined common and specific factors in client descriptions and explanations of change in structured and unstructured therapy for shyness. Clients who presented with a complaint of shyness were treated in one of three structured group therapies or an unstructured therapy group. The approach in each group was based on a different theory of change and was different in both content and focus. Content analysis
revealed that, besides specific treatment factors, all groups placed considerable emphasis on group processes and relationship factors (e.g., support). The authors suggested that the common group process factors were at least as important as the specific therapy program (p. 437).

Overall the alliance, especially as experienced by clients, is the most empirically supported common factor. Horvath and Symonds (1991) meta analysis of 20 studies between 1978 and 1990 they found an mean correlation between alliance and outcome of .26. This corresponds to an effect size of .54, which can be interpreted as saying that 7% of outcome can be explained by the alliance (2001). A more recent meta-analysis by Martin, Garske and Davis (2000) included studies through 1997 and found an alliance-outcome correlation of .22 (effect size = .45), slightly lower than that found by Horvath and Symonds. However, it was still in the medium effect size range and indicates that 5% of outcome variance is associated with the alliance. This is significant considering that Wampold’s (2001) review of meta-analyses on psychotherapy outcomes found that differences among treatments produced, at most, an effect size of .20, which indicates that about 1% of outcome was due to the type of treatment.

In giving alliance ratings clients value therapist characteristics such as therapist-provided help and demonstrated warmth, caring and emotional involvement (Bachelor, 1991, 1995; Lambert & Bergin, 1983; Lazarus, 1971; Murphy et al., 1984). This pattern also emerged in the OUHRS. The FIS measure, which includes the dimensions of warmth, acceptance, and understanding, empathy, and emotional expression, predicted higher client WAI scores at follow-up. The next common factor that is discussed,
emotional experiencing, has been shown to relate to alliance (Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996), as well as therapist provided conditions (Karon & VandenBos, 1970; Rogers et al., 1967; Schoeninger, Klein & Mathieu, 1967; van der Veen, 1965; van der Veen, 1967).

Emotional Experiencing (involvement)

Another common factor is client emotional involvement or investment in therapy. For authors of different orientations, psychotherapeutic change implies significant affective processing and learning. As part of such emotional involvement (Greenberg & Safran, 1987; Teasdale, 1993), the clients’ experiencing refers to the ability to focus on and accept their affective reactions. In a review of 1,100 quantitative studies of the relationships between process and outcome variables, Orlinsky and Howard (1986) found that “self-relatedness” was “the most consistently positive correlate of therapeutic outcome” (p. 366). Clients high in self-relatedness were defined as being “in touch with themselves…and open to their feelings” as contrasted with being “out of touch with themselves” (p.359).

Castonguay, Goldfried, Wiser, Raue and Hayes (1996) sought to examine variables that were specific to cognitive therapy and factors that this approach shares with other approaches. In this study scores on the WAI-Observer form (Horvath & Greenberg, 1986) and clients’ emotional experiencing (Klein et al., 1969) were both found to be related to improvement. In keeping with the findings of many of the studies in this thesis, the intervention focus unique to cognitive therapy was not related to client improvement.
In fact, the specific elements of cognitive therapy were positively related to depressive symptoms after therapy. A further examination of sessions found that some therapists dealt with strains in the alliance by increasing their attempts to persuade the client of the validity of the cognitive therapy rationale. In other instances, the therapist treated these strains as a manifestation of the client’s distorted thoughts, which needed to be challenged. Such interventions led to repeated cycles characterized by the therapist’s perseverance in the application of cognitive techniques and the client’s increased unresponsiveness to the treatment.

Castonguay et. al hypothesized that instead of focusing on the client’s distorted thoughts and the impact of these thoughts on emotions, the therapists may have been better off facilitating the client’s exploration of feelings, as suggested by the positive relationship between client experiencing and symptomatic improvement. This study found that therapists with high alliances tend to have higher outcomes and clients who reach higher experiencing levels. The data also seem to suggest that common interpersonal skills of attending and facilitating experiencing have an important role, especially when techniques are not going as planned. It is possible that the therapists’ interpersonal skills allow them to be flexible when rigid adherence to technique is not called for, thus allowing them to develop alliances and facilitate client experiencing. This study fits well with the results of the OUHRS which found that FIS, or basic interpersonal skills, were more important that what the techniques that the trained therapists learned in two years of training. It is possible that in the OUHRS, the trained
therapists fell back on their training techniques, rather than using their innate helping skills to work through alliance strains.

Goldman and Greenberg (2001) explored the relationship between theme-related depth of experiencing and outcome in 35 clients who received one of two types of brief experiential therapy. 17 clients received 16-20 weeks of client-centered therapy, and the 18 other clients received the same amount of process-experiential therapy. Based on therapist post-treatment interview reports, a method was developed for ascertaining the major core themes that evolved over the course of therapy. Themes tended to be confirmed by client post-session reports. Observation revealed that themes were either interpersonally or intrapersonally based. Depth of experiencing was measured at four different points over the course of therapy: one early session, and in three different, highly theme-related sessions sampled from the last half of therapy. Results indicated that a) mean theme-related depth of experiencing predicted outcome b) change in theme-related depth of experiencing over therapy predicted a decrease in depressive symptoms and increases in self-esteem, even when the alliance was taken into account, and c) level 6 and above on the experiencing scale showed the highest correlations with outcome. This study showed that change in the depth of experiencing over the course of therapy is related to outcome, suggesting that increased depth of experiencing is related to process change. Early session experiencing also predicted outcomes. Thus, experiencing may be both a client trait variable, indicating an ability to look inwards, and a change process.

Wiser and Goldfried (1998) examined the effects of different interventions on experiencing in sessions of psychodynamic-interpersonal therapy obtained from
experienced clinicians in a naturalistic setting. Reflections and acknowledgments, affiliative and noncontrolling interventions, or interventions highlighting nonspecific client content were associated with maintained high experiencing. Lengthier interventions and interventions rated as affiliative but moderately controlling were associated with shifts to low experiencing. While it is difficult to extrapolate these findings towards a model of therapist skill, one could hypothesize that high skilled therapists would deliver interventions that are more affiliative and less controlling.

Watson and Greenberg (1996) identified a pathway from in-session process, and problem resolution, to post-session change and final outcome. Two brief treatments for depression, one using client-centered and the other process-experiential interventions, were compared on client process and outcome. The process-experiential group showed significantly higher levels of experiencing, vocal quality and expressive stance, and greater problem resolution than the client-centered group in 2 of 3 PE interventions studied. Clients’ degree of problem resolution correlated significantly with depth of experiencing, and sustained resolution over treatment resulted in better outcome.

In the first investigation concerned with the evaluation of changes in expressed experiencing, Gendlin, Jenney, and Shein (1960) looked for an association between outcome and the patient’s use of the relationship for significant experiencing versus a simple concern with the relationship as a topic. They asked 16 therapists to rate 39 of their neurotic cases at the seventh and final interview on six process dimensions, several of which related directly to experiencing. Comparisons of the ratings with independent evaluations of case outcome showed that during therapy the more successful patients
moved from remote talk about feelings to direct expression of feelings. They became more able to see the relationship with the therapist as an important new experience, and had worked with parallels between experiences in the therapy context and problems experienced in other contexts. These findings pointed to experiential involvement and shifts as factors in successful outcome.

The most extensive early application of the experiencing scale was the study of 14 schizophrenic patients in therapy matched to 14 control patients not in therapy, carried out at the University of Wisconsin by the Psychotherapy Research Group (Rogers et al., 1967). Experiencing ratings were made by four judges for a large number of segments taken from therapy interviews and non-therapy sampling interviews. Experiencing level, both as averaged across the first 30 sessions and as averaged over the entire course of therapy, was related to independent judges’ ratings of therapist accurate empathy, and to the patients’ perceptions of favorable therapists’ attitudes on the Relationship Inventory. Mean experiencing level for the treatment course was associated with some indices of change and outcome, most strongly with MMPI Sc scale improvement and to a lesser extent with MMPI Hs and Pd scale improvement, self-rated (Q-sort) adjustment, clinicians’ evaluations of change, and percent of time out of the hospital. In general, the relationships between experiencing and outcome revealed by this study affirmed the value of the experiencing construct and the validity of the scale. The study showed most clearly that the scale taps differences in patients’ verbal behavior, which is a function of initial expressiveness, capacity of motivation for therapy, and therapist empathy and genuineness in therapy, which is predictive of a good outcome.
Experiencing was also used along with other processes and conditions scales (including Relationship, Problem Expression, Personal Constructs; Empathy, Warmth, and Congruence) in a study of the effectiveness of psychoanalytic psychotherapy with schizophrenics carried out by Karon and VandenBos (1970) as part of the Michigan State Psychotherapy Research Project. In this study 36 patients were randomly assigned either to interpretative psychoanalytic psychotherapy without drugs, to ego-analytic psychotherapy with drugs, or to supportive therapy with drugs. Relationships between experiencing and the therapist variables were complex. Experiencing was positively related to therapist warmth. It appears that therapist consistency affects patients in a different way. The more consistent the therapist was, the higher the client’s mean experiencing. While therapist consistency raised mean client experiencing, it also made the client’s experiencing more inconsistent. The more consistent therapists had clients with greater mode-peak difference scores. This finding hints that the therapist who can offer a stable relationship to the client can induce the client to vary his or her experiencing, presumably to higher levels of the scale. Experiencing mode, peak and difference scores were very strongly related to ratings of health made from the clinical status interview and to the length of hospitalization.

van der Veen (1965; 1967) used two of the process strands, experiencing and Problem Expression, to study the relative effect of the patient and therapist on one another’s behavior. His study showed that process level and process improvements over therapy were moderately related to the quality of the therapist’s performance. Therapist’s behavior was rated on the Congruence (genuineness) and Accurate Empathy scales.
There were consistent, significant positive correlations between process level and therapist genuineness and empathy rated in each unit, and averaged over both interviews, and between improvement by the patient in experiencing and the therapist provided conditions level over the interview. Generally, van der Veen found that therapist conditions and performance were related to client experiencing.

Schoeninger et. al (1967) used the experiencing scale to study several variables in the client-therapist interaction, specifically the effect of training designed to enhance client experiencing, and the effect of therapist self-disclosure in sessions. There were two experimental dimensions: therapist self-disclosure and non-disclosure and either the presence or absence of training. Subjects were 32 students motivated by their desire to discuss a problem. After three 60 minute sessions, clients rated their therapists on certain variables including comfort, genuineness, and empathy. The hypotheses about the effects of specific therapeutic practices on experiencing revealed few significant correlations, primarily because of the unanticipated differences in the therapists’ performances of their roles. During the brief series of interviews, experiencing level was closely related to the patient’s characteristics and expectations, but did not reflect the attempted manipulation of client’s interview behavior through pre-training. Experiencing levels did not respond to the technical variations in the therapist’s behavior but did correlate with therapist qualities as perceived by the clients.

This particular study mirrors many of the larger trends found in psychotherapy research discussed in this paper. The specific therapist intervention of self-disclosure, as well as therapist training to elicit client experiencing, had little effect on clients’
experiencing. The therapists varied greatly in their ability to accurately perform the technical elements demanded by the study. Finally, the more general interpersonal skills did influence experiencing.

Summary of Common Factors

Common factors have shown a significant contribution to outcome. A single common factor, the therapeutic alliance accounts for 5% of outcome variance (Wampold, 2001). Could the therapist’s ability to provide common factors or their general skills be what determines outcome? Given that the ability to provide common factors is linked with general skills, not learned techniques, it seems likely that there would be a normal distribution of the ability to provide these common factors. Therefore, we could assume that on average, both groups of therapists in an RCT would have a similar amount of the ability to provide and utilize common factors, thus helping to explain the Dodo verdict. This varying ability in providing common factors would also explain differences in therapist performance. As mentioned previously, therapist differences frequently outshine treatment differences (Garfield, Affleck, & Muffly, 1963; Luborsky et al., 1986; Orlinsky & Howard, 1986). In studies that have sought to tease apart therapist differences, it the ability to provide and leverage common factors that differentiates the therapists. The body of literature which reveals therapist effects, as well as studies that seek to tease apart therapist differences, will now be reviewed.
Therapist Effects

Therapist effects refer to the degree to which therapists vary in the outcomes they produce, apart from the effects due to treatments (Wampold, 2001). As mentioned earlier in the Dodo Bird section, treatment differences are often found to be smaller than differences attributed to therapists (Beutler et al., 1994; Crits-Christoph et al., 1991; Lafferty et al., 1989; Orlinsky et al., 1994). Despite the strict controls used in the NIMH Treatment of Depression Collaborative Research Program (TDCRP), which was the largest and most rigorous RCT attempted to date, therapist effects were evident. The study compared the effectiveness of CBT, IPT, and drug therapies (Elkin, 1994). The researchers attempted to treat therapists as error variance. Even though the therapists possessed equivalent levels of prior professional experience and were systematically instructed to use criteria based methods over a two-year period, therapists varied widely both in the levels of specific therapeutic skills achieved and in their rates of effectiveness with clients. In a subsequent reanalysis of the NIMH study, Blatt (1996) demonstrated a wide range of mean improvements within treatment modalities for the 28 therapists studied. A mean therapeutic effectiveness measure was derived for each of the 28 therapists based on the aggregate of residualized therapeutic change scores of the 5 primary outcome measures for each patient at termination. The distribution of the therapists was divided into thirds, and comparisons indicated that more effective therapists were more psychologically minded, eschewed biological interventions (i.e., medication and electroconvulsive therapy) in their ordinary clinical practice, and
expected outpatient treatment of depression to take longer than moderately effective and less effective therapists.

Therapists effects were also found in the Sheffield Psychotherapy Project, which compared CBT to interpersonal-psychodynamic treatment (Shapiro & Firth-Cozens, 1987). Four therapists each saw 6 to 18 clients in a crossover design intended to hold constant individual differences of clients and therapists, while maximizing treatment differences. Even though all four therapists were following the same treatment manual, one of the therapists had much better results and was responsible for most of the success obtained (Shapiro, Firth-Cozens, & Stiles, 1989). This demonstrates, as reported by Henry et. al (1993) and Beutler (1997), that systematic, manual guided training does not eliminate the presence of significant variability among the individual therapists’ rates of efficacy.

Luborsky et. al (1986) reanalyzed four studies to determine the size of therapist effects. They obtained the data for the Hopkins Psychotherapy Project (Hoehn-Saric et al., 1964), The VA-Psychotherapy Project (Woody et al., 1983), The Pittsburgh Psychotherapy Project (Pilkonis, Imber, Lewis, & Rubinsky, 1984), and the McGill Psychotherapy Project (Piper, Debbane, Bienvenu, & Garant, 1984). Luborsky et. al considered the therapists as a random factor as recommended by Crits-Christoph and Mintz (1991). They found no differences among treatments being compared but a great variation in therapist effectiveness. The authors concluded that the size of the therapists’ effects generally overshadowed any differences between different forms of treatment in those investigations.
Lafferty et al. (1989) examined differences between more and less effective trainee psychotherapists. They found that one third of therapists who were working with outpatient clients produced as many negative therapeutic effects as positive ones, independent of the type of therapy utilized. In contrast, other therapists in this study were uniformly effective with most patients. Lambert (1989) reviewed outcome studies and developed a model which stated that the largest amount of variation in therapeutic outcome was attributable to differences in clients’ characteristics, with the second largest amount accounted for by therapist differences. Variations in theory and technique came in a distant third. Thus, according to this model, RCT designs attempt to eliminate the two most important sources of variance, client and therapist variability. Despite this, therapeutic techniques and their underlying theoretical explanations of change continue to be emphasized in graduate training, research grants, and textbooks, while minimal attention is given to the therapist’s personal qualities and their impact on outcome (Lambert & Okishi, 1997).

Therapist effects also have implications on the design and statistical analyses of studies. Christoph and Mintz (1991) found that researchers have often made the mistake of treating therapists as equal. They reviewed many past studies in which therapist performance was in fact not equal. They recommend that therapists should be treated as a random factor, indicating that they are being sampled. Failure to do this results in an unacceptably high risk of type I errors and could explain many findings of treatment differences. As mentioned earlier, Wampold and Serlin (2000) also examined the effect of improperly modeling nested designs in psychotherapy studies. They believe that if
therapist effects are properly modeled, the effect size of .20 for treatments is completely artifactual. They conclude that in reality treatment effects appear to be zero.

Exploring Therapist Effects

Several researchers are addressing the uniformity myth and devoting resources to studying the contributions of therapists to outcome. There is considerable support for the view that the individual therapist’s attributes, attitudes, and actions (e.g., interpersonal skills, countertransference propensities, and personality) match or override the effects of particular techniques, (Beutler et al., 1994; Crits-Christoph et al., 1991; Lafferty et al., 1989; Orlinsky et al., 1994). Examining how these factors lead to change is one of the major challenges of the field. Lambert (1995) argues for research strategies that measures the outcome of each individual therapist’s caseload. Researchers can look at how each therapist did with their caseload and then examine the therapy and therapist qualities associated with that performance to see how therapist variables are related to treatment outcome. Direct assessment of skill level has considerable face validity and circumvents some of the confounding variables of age, training, experience, and orientation (Beutler et al., 1994). By linking skills and therapeutic operations to productive therapy moments and outcome, researchers can see which skills and operations are most salient and which are simply emphasized out of tradition.

Common Factors Help Explain Therapist Success

When researchers have found evidence which could possibly explain the case of differential therapist effectiveness, it has usually been the general or common factors that
distinguish between those who are more or less successful. Luborsky et. al (1985)
reviewed what contributed to varying success levels of nine therapists who used
supportive-expressive psychotherapy, cognitive-behavioral psychotherapy, or drug
counseling approaches. They found that patient characteristics within each caseload were
similar and could not explain differences. Therapist personal characteristics were
correlated with outcome, but not significantly. An early measure of the alliance as
measured by the Helping Alliance Questionnaire provided significant correlations (.65)
with outcome. Among therapy techniques, “purity” provided significant correlations (.44)
with outcome. Purity was determined by raters who assessed the degree to which the
therapist adhered to the manual for the treatment approach. The authors concluded, “the
therapist’s ability to form an alliance is possibly the most crucial determinant of his or
her effectiveness.” In trying to explain the relationship of purity of technique to outcome,
they posited that once a helping alliance is established, the therapists who do what they
are supposed to do (provide pure therapy) achieve their effectiveness in this way. Another
explanation is that when the therapist and client enjoy a strong alliance, the client enables
the therapist to adhere to his or her intended technique.

Luborsky (1997) studied 22 therapists’ caseloads with drug-addicted and depressed
clients and found important differences in the post-treatment outcomes among the
therapists. The patients were seeing similar clients, and therapists were selected for their
competence in their particular form of therapy. Despite efforts to minimize differences
between the therapists, the range of improvements for the clients of the 22 therapists
varied widely from a slightly negative therapeutic impact to more than 80%
improvement. Therapists’ qualities, as judged by their peers, showed that interest in helping patients (.44) and the therapists’ psychological health and skill (.41) correlated best with client outcome. In addition, therapists whose caseloads showed the most improvement were those who were able to engage and retain the largest proportion of assigned patients suggesting a greater ability to establish a working alliance and provide clients with the reparative experiences they needed. Although there were differences for each therapist within their individual caseload, it was the therapists’ capacities or skills that were responsible for change rather than differences in clients’ characteristics or diagnoses. Thus, better therapists did better with most of the clients they treated. However, the two best therapists in the study were very different. One used mostly supportive techniques, whereas the other used mostly expressive techniques. Thus, there are different routes to success and it may not be wise to always make the specific type of therapy the object of the research. Even though the technical routes differed for these two highly effective therapists, they both showed one common trait - a strong working alliance with their clients.

In a study by Najavits and Strupp (1994) 16 therapists who had been recommended by their supervisors as “caring empathetic clinicians” were divided into two groups based on the success of their clients on objective outcome measures. Although training and specific technical skills were not related to outcome, nonspecific qualities did distinguish the two groups. As a group, the more effective therapists demonstrated significantly more warmth and understanding, and they had fewer recorded instances of ignoring, neglecting, attacking, and rejecting their clients than less effective therapists. The authors
suggested that “basic capacities of human relating-warmth, affirmation, and a minimum of attack and blame—may be at the center of effective psychotherapeutic intervention. Theoretically based technical interventions were not nearly as often significant in this study” (p. 121). This was similar to the pattern that emerged in Strupp and Hadley’s (1979) earlier study comparing licensed therapists to college professors.

Thus, therapists with better outcomes are often the ones who have the general skills to relate in a warm, empathic manner and build strong alliances. The OUHRS study sought to further explore the contributions of therapist interpersonal skills. The OUHRS took measures of therapists interpersonal skill (FIS) to determine its influence on outcome. Further, it sought to determine if FIS would be more influential than level of training. The results of the OUHRS study will be reviewed, and suggestions for how to build upon its findings will be proposed.

The OUHRS

An ideal way to examine the contributions of specific versus common skills would be to design a study that included therapists who varied across both dimensions. The OUHRS study compared psychotherapy outcomes of trained (2+ years of graduate study) and untrained (graduate students from other disciplines) therapists who also differed in their level of Facilitative Interpersonal Skills (FIS). The FIS rating was made based on the Social Skills Inventory and the Facilitative Interpersonal Skills Test. This provides the two by two design seen in Table 1.
The results indicated that there was no difference by level of training on outcome measures. This mirrors the results from the majority of studies, which have examined the effects of experience on training. In the OURS there was a surprising trend ($p = .06$) towards the untrained having *higher* alliance measures. As expected, high FIS therapists had higher alliances and higher outcomes. Thus, the pre-study evaluation of FIS predicted higher outcomes. It seems therapists used their FIS to develop better working alliances. There was a strong interaction between client social skill as measured on the SSI and therapist FIS. The dyads that had the best outcomes were the ones in which both participants had high social skill.
The Current Study: Linking Skill to Process

Given the interesting results in the study, a question begging to be answered is what process events were related to the therapist Facilitative Interpersonal Skills (FIS)? How did FIS influence therapy? FIS led to higher alliance ratings, but it remains unclear how FIS influenced the moment-by-moment observable process within sessions.

This study presented the opportunity to search for a common process element that may be influenced by therapist skill and contribute to outcome. As previously described, therapists qualities such as warmth, empathy and understanding contribute to outcome. The therapeutic alliance is one of the strongest predictors of outcome.

This thesis called for a process measure that was free of ties to specific techniques, had demonstrated links to outcome and was influenced by therapist interpersonal skills. Experiencing was chosen as the primary measure of this study because it met these criteria. Experiencing or emotional involvement of the client is included in most models of common factors. The experiencing scale is free from ties to any specific techniques and therefore fits a study which was designed to examine the contributions of common versus specific skills. Higher experiencing has also been shown to be related to better therapeutic outcomes. Experiencing has been shown to be influenced by therapist influenced common factors such as empathy, warmth and the alliance and will therefore, hopefully reveal a process difference between high and low FIS therapists.
Hypotheses

All hypotheses for the present study involve the client experiencing scale, which consists of a mean and mode rating for each segment (See Measures section for detailed description). The mean for the mode scores (mean-mode) and the highest mode score (peak-mode) were used as variables, as well as the corresponding mean-peak and peak-peak scores. The first two hypotheses relate to the construct validity of the experiencing scale within the present sample. Hypotheses 3 is an overall test of the experiencing predictive validity. Hypotheses 4 and 5 involve more specific tests of predictive validity that relate to the design of the OUHRS.

Hypotheses:

1) Therapeutic processes: The alliance

There is a positive relationship between mean client-rated WAI scores and mean-mode and mean-peak experiencing scores over therapy.

2) Therapeutic processes: narrative process

There is a positive relationship between session narratives that are more emotionally internal and meaningful and mean-mode and mean-peak experiencing scores, and there is a negative relationship between narratives that are more external and mean-mode and mean-peak experiencing scores.

3) Outcome

There is a positive relationship between outcome, as measured by the IIP and OQ residualized change scores at follow-up, and peak-peak experiencing scores.

4) Therapist FIS
High FIS therapists will have clients with higher mean-mode, peak-mode, mean-peak and peak-peak experiencing scores than low FIS therapists. This hypothesis is tested using two MANOVAs with therapist FIS, client social skill, and training as the fixed factors. First the mean-mode and peak-mode are entered as the dependent variables followed by the mean-peak and peak-peak scores. Further explanation and rationale for this appears in the Results section.

5) Client social skill

Clients with high social skills will have higher mean-mode, peak-mode, mean-peak and peak-peak levels of experiencing than low social skill clients. This hypothesis will also be tested using two MANOVAs with therapist FIS, client social skill, and training as the fixed factors. First the mean-mode and peak-mode were entered as the dependent variables followed by the mean-peak and peak-peak scores.

METHOD

Sample

Participants. Forty-five undergraduate students enrolled in introductory psychology courses at a state university in Ohio participated in OUHRS study, which was conducted prior to the current study. The students who qualified for the study showed significant levels of psychological distress and adjustment problems. Twenty-seven female and eighteen male participants completed all phases of the study. Ages ranged from 17 to 23 with a mean of 19. Thirty-six percent reported some prior psychotherapy experience. 90.9% of participants were Caucasian, 2.3% were African-American, 2.3% were
Asian/Pacific Islander, and 4.5% were Multiracial. The current study used all participants.

Participation was voluntary, and participants earned four experimental credits and fifty dollars for completion of the study. Participants who were randomly assigned to the therapy condition were asked to commit 13.5 hours to the study over the course of 22 weeks. Control participants were awarded $10 for four hours of participation over a 22 week period. Students were assured that all information would be kept confidential. In addition participants were informed that they could drop out of the study at any time without penalty if they were uncomfortable with the study or wished to seek psychotherapy elsewhere.

**Therapists.** Eight of the therapists were male and fifteen were female. The eleven trained therapists came from the clinical psychology Ph.D. program at Ohio University. All trained therapists had completed at least two years of the program during which time they took classes in basic helping skills, psychopathology, personality, courses in interpersonal and cognitive behavioral interventions, a survey of the treatment literature, and ethics. During their second year, these students saw clients through practica which were supervised by professors. The clients during the second year were usually referred from the student health center. The twelve untrained therapists came from a variety of other graduate programs at Ohio University including: 3 from Biology, 1 from Chemistry, 2 from Experimental Psychology, 1 from History, 1 from English, 1 from Comparative Arts, 2 from Interpersonal Communication, and 1 from Consumer Sciences. Therapist ages ranged from 23 to 53 with a mean of 31. The clinical psychology
therapists included 9 Caucasians, 1 Hispanic, and 1 Asian. The untrained therapists were 10 Caucasians and 2 Asians. Therapists from the psychology department and other disciplines were paid $200 for their participation. They committed 11 hours over a seven week period to meet with undergraduate participants and complete questionnaires. Therapists were assured that all of their questionnaire and recorded session data would be held in confidence and their identity would be protected by assigned identification codes.

Patient Selection and Participation. 2142 undergraduate students who enrolled in introductory psychology courses between March 1998 and March 1999 were screened for general psychological distress and depression using the SCL-90-R. Undergraduate students who scored at least two standard deviations above the mean were invited to participate in a more in-depth screening. This involved an assessment interview with a licensed clinical psychologist or one of five advanced level clinical psychology doctoral candidates with at least 1,000 hours of applied clinical experience. Students were given a second SCL-90-R and the Inventory of Interpersonal Problems-CX (Horowitz, Rosenberg, Baer, 1988) which measures social competence and interpersonal skills.

Two hundred four participants qualified to continue in the study. In order to qualify both SCL-90-R Global Index Severity Index scores needed to be two standard deviations above the mean, at least three subscales of the IIP needed to be clinically elevated (scores above 10), and upon completion of the assessment interview they had to have received a DSM-IV diagnosis which interfered with interpersonal functioning. Students presenting with DSM-IV diagnosable alcohol/drug abuse or dependence, frequent suicidal ideation or severe personality disorders were excluded from the study and referred to counseling.
Students currently receiving psychotherapy, or those who planned to enter psychotherapy in the next 22 weeks were also excluded.

Of those screened students who were eligible to complete the study, 83 agreed to participate in the study. Participants were randomly divided into “active” (59 participants) and “control” (24 participants) groups. Active participants were asked to meet for seven weekly one-hour sessions with one of the “therapists” to discuss their problems and fill out questionnaires. They were also asked to complete termination and follow-up questionnaires and an interview with one of the assessment clinicians. Control participants were asked to complete a termination interview, and questionnaires on three separate occasions over a 22 week period. 52 active participants completed the study and 7 dropped out. The active participants’ data was included as complete if at least 5 sessions and termination had been completed. Only data from “actives” were used in this study.

**Therapist selection and participation.** Fifty-six graduate students from a variety of disciplines including psychology responded to advertisements regarding participation in this study. All completed a social and interpersonal skill questionnaire, the Social Skills Inventory (Riggio, 1989) and those scoring in the high and low ranges were asked to complete a performance analysis.

**Performance Analysis.** Two separate performance analyses were conducted. In the first performance analysis, the therapist viewed a five minute segment of a videotaped exchange between a client and a therapist. Immediately after viewing the tape, the therapist was asked to talk about his/her perceptions of the nature of the relationship and
the feelings both the therapist and the patient might have been experiencing. This aspect of the performance analysis was modeled on an analogue developed by Bein and Strupp (1996).

The second part required therapists to demonstrate the skills they would use to respond to videotape examples. These videotape examples consisted of actors reenacting different problematic situation encountered in the Vanderbilt II study. All of these situations might threaten the formation of a therapeutic bond and were deemed to be challenging for the therapist. The first segment involved the actor playing the role of a self-focused, highly negative, and self-effacing client. The second was a ten-minute encounter with an actor playing the part of an other-focused, friendly, but highly dependent client who frequently asked for advice from the therapist. These scenarios were chosen because they provided the therapist with interpersonal challenges from opposite ends of the interpersonal circumplex model and because they were situations that are often difficult for therapists to address (Benjamin, 1993).

Videotapes of the interview and performance analysis were rated by two doctoral level psychotherapy researchers. Each rater was asked to place the therapist participants into the high or low skill group based on their interpersonal perception and ability to take interpersonal actions (i.e., performance analyses data). Raters were provided with a rating scale and sample anchor ratings derived from pilot testing of the performance analysis procedure.

Therapists who were below the median on the interpersonal competence questionnaire (i.e., SSI) and who were rated as having low interpersonal skills from the two
performance analyses were assigned to the low interpersonal skills group. Likewise, therapists above the mean on the questionnaire and who were judged to have high interpersonal skills from the performance analyses were assigned to the high FIS group. Therapists who did not clearly fall into one of these groups were assigned to one only if their score on interpersonal competence was within one standard deviation of the group. Two therapists who were between groups were placed in the low group according to the above mentioned guidelines. Each therapist was assigned to two “active” client participants.

Procedure

Weekly Sessions. All therapy sessions took place at a university psychology training clinic. Clients met with therapists on a weekly basis for seven sessions. All therapists were instructed about the general parameters of typical therapy sessions (e.g., 45-50 minute sessions, meetings take place in assigned clinic rooms). Therapists were also instructed about expectations for ethical conduct and received a copy of the ethical guidelines published by the American Psychological Association (APA, 1992). With these basic constraints in mind, therapists were asked to help students with their problems in any way they believed would be helpful and respectful. No advice was offered to therapists as to the use of specific “techniques”. Therapists met periodically with the principal investigator, a licensed clinical psychologist, in order to process any questionable situations or interactions. Further, therapists were asked to directly contact the principle investigator to discuss any situations they felt unable to address. All sessions
were recorded on audiotape and closely monitored to assure the safe and ethical conduct of the sessions. Clients were informed about confidentiality, the taping procedures and other treatment options prior to the beginning of therapy. They were also informed that their “therapist” was a graduate student who may have had little or no formal training in therapeutic techniques.

Audio Tape Sample. For the current study, which builds upon the OUHRS study, eight-minute segments were sampled (30th to 38th minute) from sessions 2, 3, 4 and 6 for coding. Thus the raters would be giving one mode and one peak rating to the entire eight-minute segment. The segment of the 30th to the 38th minute was selected because in an early experiencing study (Kiesler, Klein, & Mathieu, 1965) which compared 5 successive 8-minute portions of one therapy hour for each of 24 patients (8 schizophrenics, 8 neurotics, and 8 normals), differences in experiencing between the groups were most evident towards then end of the hour. The Klein manual states that is up to researcher to select the segment that meet their researcher’s needs. The eight minute length was chosen because the experiencing scale was designed for use with tape recordings and transcripts for units of 2-8 minutes in length (Klein, Mathieu-Coughlan, & Kiesler, 1986). Kiesler et. al (1964) conducted a study to compare the effects of coding different length segments. Scale ratings were made for 2, 4, 8, and 16-minute segments drawn from each of two therapy hours with eight normal, eight neurotic, and eight schizophrenic subjects. Only the absolute experiencing level proved to be influenced by segment length. However, the reliabilities, range, and discriminatory power of the ratings were independent of the length of the segment. Klein et. al (1986) advise that segments of 5 to 8 minutes provide
the rater with enough information to identify high levels of experiencing without becoming unwieldy or tedious. In a study by Mintz experiencing and other process ratings of a 4-minute segments were found to be highly correlated with ratings of the hour as a whole, indicating the adequacy of brief sampling (Mintz, 1969).

Studies using other psychotherapy scales have also shown the adequacy of brief sampling. Weiss et. al (1988) found no significant differences when rating entire therapy sessions versus the first half of a session using the Therapist Action Checklist/Patient Action Checklist (Hoyt, Marmar, Horowtiz, & Alvararez, 1981). In their review of the Vanderbilt Psychotherapy Process Scale (VPSS), Suh, Strupp and O’Malley (1986) found that 10-minute segments (Gomes-Schwartz, 1978), 15-minute segments (O'Malley, Suh, & Strupp, 1983), and then entire therapy hour (Gomes-Schwartz & Schwartz, 1978) demonstrated significant relationships to outcome, which they state may provide indirect support for the representativeness of segments. Based on these findings, the authors suggest that 15 minute segments may be adequate for capturing the process characteristics that the VPPS is designed to assess.

These studies which reveal the representativeness of brief segments also provide some evidence that random experiencing samples may not always be advantageous. Klein et. al (1969) state that in studies where the focus is on the characterization of a fairly homogeneous patient group, sampling a single time period within the hour may well control for the confounding effect of time, but may not reflect the highest levels or full range of experiencing operative in that group. They go on to state that sampling of standard time periods may have more precision and offer better control over the time
factor in small samples, but it may not always yield representative data or data that can be readily interpreted without supporting information regarding the specific nature of intra-interview process trends. Data on the representativeness of the 30-38th minute segment used in this thesis were presented as a possible limitation in the Discussion section.

Other researchers have used single time periods rather than random time periods. Castonguay et. al (1996) elected to use only one session but coded segments at the beginning, end, and middle of each hour. They did not use a randomized approach, but coded the same time points for every hour. Rennie, Brewster, & Toukmanian (1985) coded the experiencing of therapists in training who were acting as clients for 20 minute interviews, to determine if their openness to and expressiveness of inner experiencing would relate to their acquisition of counseling skills. They selected two five-minute segments, the forth to the ninth minute and the twelfth to the seventeenth minute.

Lansford and Bordin (1983) attempted to compare the experiencing and free association scales. They coded a smaller number of tapes, 30, and coded the first four minutes, the middle two minutes and the last four minutes of each tape.

Other psychotherapy scales have also been used on non-random segments of sessions. In attempting to validate the Penn Helping Alliance Scale, Alexander and Luborsky (1986) coded the first 20 minutes of psychotherapy sessions. They selected 20 minutes rather than the more typical five-minute segments used in psychotherapy research because they felt five minutes was too short for judging a relationship variable such as the alliance (Mintz & Luborsky, 1971). O’Malley et. al (1983) used segments consisting
of a composite of the first, middle, and last five minutes of therapy sessions with the VPPS.

Four sessions were chosen because it is important to sample from at least two periods in therapy if experiencing is to be considered in relation to outcome criteria (Klein et al., 1986). Some of the past studies suggest that it is also desirable to include samples from the midpoint or working phase of therapy (Berman & Norton, 1985). Sampling only initial or terminal hours risks missing the patient’s actual work and progress and getting instead only routine arrangements, history taking, or termination material involving future plans, external details, a review of progress, or a period when defenses are deliberately re-instituted. The more samples that are taken from each case, the greater is the likelihood of getting reliable estimates of the variability, range and trends in experiencing (Klein et al., 1986).

Several tapes were missing from the archives. In this case an adjacent tape was substituted. If a session two tape was missing, a session one tape was substituted, as session 3 was already to be included in the data. If a session 4 tape was missing, a session 5 tape was used. If a session 6 tape was missing, a coin flip determined if a session 5 or 6 tape would be substituted. However, sometimes more than one tape would be missing for a given therapist. Possibly certain therapists were not diligent about recording their sessions. In this case the closest available session in terms of session numbers would be used. Four missing session 2 tapes were replaced with two session 1 tapes and a session 5 tape. Two missing session 3 tapes were replaced with session 1 tapes. Three missing session 4 tapes were replaced with two session 5 tapes and a session 7 tape. Six missing
session 6 tapes were replaced with four session 5 tapes and two session 7 tapes. In total fourteen tapes were replaced with tapes from the same dyad.

Audio Tape Preparation. Random numbers were assigned to the therapy sessions so that the coders would not know which dyad they were coding or the session number. These numbers were affixed to the cassettes using removable labels which covered the true session information. All sessions in the OUHRS study were intended to be 50 minutes long. Coders found this segment by using the counter on top of the cassette player. Using a stopwatch the first author found the counter numbers that pertain to the 30th and 38th minute. These counter numbers were affixed to each tape player for easy guidance. The coders were instructed to rewind the tape to the beginning, set the counter to zero and then fast forward to the counter number pertaining to the 30th minute. They coded until they reached the counter number pertaining to the 38th minute.

Tapes were assigned in a random order, using a Microsoft excel function which generates a random number between 0 and 1. An Excel formula gave each coder a 25% chance of being assigned to that tape based on the random number that was generated. In order to prevent coders from becoming involved in the “story” of a particular therapist-client dyad, care was taken not to assign one coder more than two sessions for a particular dyad.

Assignment Sheets. Coders were given assignment sheets listing all of the tapes which they were to code. This sheet contained spaces for the tape id, the date coded, coding time, coding difficulty (1-5 scale), and coder comments. Coders were encouraged to place questions about a particular tape in this comment box or explain their rationale
for giving a certain rating. This information was helpful during consensus coding. The first author entered the codes from the assignment sheets into an excel spreadsheet on a weekly basis.

**Coder Selection and Training.** Four undergraduate coders were selected from six who interviewed for positions with the project. They were given course credit for their work. The training techniques for the scale have been constructed so that nonprofessional, clinically naive people can serve as raters. Nonprofessionals may be freer from biases concerning patient types, therapeutic techniques, or research goals (Klein et al., 1969). Coders were introduced to the project and asked to sign a confidentiality agreement. Coders were trained using the training manual and materials provided by Klein et al. This manual consists of audio-taped segments that the coders were to rate on their own. The coders then met together with the first author to compare their scores and discuss the rationale for their scores. The first author, who was already familiar with the system, went through the first three weeks of training with the coders. After that he did not code, except to participate in consensus meetings. Reliabilities were tested at every stage. Table 2 shows the reliabilities for weeks 1-3. Table 3 shows reliabilities for the final assignments which the manual states are more difficult.
Table 2
Correlations between each coder and Klein manual weeks 1-3

<table>
<thead>
<tr>
<th>Score</th>
<th>Dave</th>
<th>J</th>
<th>D</th>
<th>A</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>.86</td>
<td>.78</td>
<td>.78</td>
<td>.91</td>
<td>.71</td>
</tr>
<tr>
<td>Peak</td>
<td>.82</td>
<td>.82</td>
<td>.73</td>
<td>.89</td>
<td>.74</td>
</tr>
</tbody>
</table>

* Coders are identified by letter only, except for the first author.

Table 3
Correlations between each coder and Klein manual final assignments

<table>
<thead>
<tr>
<th>Score</th>
<th>J</th>
<th>D</th>
<th>A</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>.79</td>
<td>.71</td>
<td>.74</td>
<td>.75</td>
</tr>
<tr>
<td>Peak</td>
<td>.72</td>
<td>.72</td>
<td>.76</td>
<td>.72</td>
</tr>
</tbody>
</table>

Because the coders met the acceptable reliability, they began coding the tapes. After the coders were immersed in the OUHRS tapes their reliability was checked every week.
Forty percent of tapes were double-coded to assess reliability, and these tapes were checked each week. A similar percentage, 25%, was used in a study to validate the McMaster Family Therapist Coding System (Pinsof, 1986). All coder codes were entered into an Excel spreadsheet that automatically recalculated the correlation between all pairs of double coded tapes. Thus, different combinations of coders were contributing to this overall reliability number. This correlation never dipped below .70 for mode or peak. At the end of the project the reliability was .731 for mode and .792 for peak ratings.

After the project the reliability of each coder combination was checked. Table 4 shows the correlations between coders on mode and peak scores. It appeared that coder L had some of the lowest correlations and correlations that were below the acceptable limit. Correlations comparing each coder to all other coders were also made (See table 5). These also revealed that L had lower correlations when compared to all other coders. Therefore, the tapes coded by L were reassigned to the other 3 coders (See table 6 for reliabilities after recoding all of L’s tapes).
Table 4

Correlations between coders for mode and peak scores

<table>
<thead>
<tr>
<th>Coder</th>
<th>Mode</th>
<th>Peak</th>
<th>Mode</th>
<th>Peak</th>
<th>Mode</th>
<th>Peak</th>
<th>Mode</th>
<th>Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. J</td>
<td>-</td>
<td>-</td>
<td>.78</td>
<td>.84</td>
<td>.42</td>
<td>.61</td>
<td>.75</td>
<td>.77</td>
</tr>
<tr>
<td>2. D</td>
<td>-</td>
<td>-</td>
<td>.72</td>
<td>.74</td>
<td>.85</td>
<td>.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. L</td>
<td>-</td>
<td>-</td>
<td>.63</td>
<td>.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. A</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5

Correlations between each coder and all other coders

<table>
<thead>
<tr>
<th>Coder</th>
<th>Mode</th>
<th>Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A</td>
<td>.80</td>
<td>.83</td>
</tr>
<tr>
<td>2. D</td>
<td>.75</td>
<td>.84</td>
</tr>
<tr>
<td>3. J</td>
<td>.73</td>
<td>.78</td>
</tr>
<tr>
<td>4. L</td>
<td>.61</td>
<td>.70</td>
</tr>
</tbody>
</table>
Table 6

Correlations between coders after removal of coder L

<table>
<thead>
<tr>
<th>Coder</th>
<th>Mode</th>
<th>Peak</th>
<th>Mode</th>
<th>Peak</th>
<th>Mode</th>
<th>Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. J</td>
<td>-</td>
<td>-</td>
<td>.84</td>
<td>.79</td>
<td>.81</td>
<td>.82</td>
</tr>
<tr>
<td>2. D</td>
<td>-</td>
<td>-</td>
<td>.82</td>
<td>.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Consensus Meetings. Weekly one-hour consensus meetings were held in order to prevent coder drift. The first author selected tapes on which the coders had the biggest ratings discrepancies. He also selected tapes so that all the coders had chances to have their tapes reviewed. This enabled each coder to benefit from the group discussion about his or her coding. These meetings began with a general discussion of any problems the coders were having with the coding system. Coders also had the opportunity to discuss any emotional reactions to the tapes. Given that the coders signed a confidentiality agreement, the consensus meetings were the only opportunity to process any emotional reactions which the tapes evoked. If any problems arose, coders would have been referred to trained mental health professionals. The coders did not have any severe emotional reactions.
Each segment selected for consensus coding was listened to in its entirety. The coders were not told which tape had been selected, nor what scores it was given. Two of the coders would have coded the tape before, but they may not have remembered what scores they gave it. In this manner, all the coders had to listen and make their own attempt at coding the tape. Afterwards, all coders were asked to give their scores and rationale. The first author then reviewed the actual scores that were given by the two original coders. The group used the manual to apply the rules of the coding system until a consensus about the scores was reached.

When particular coding questions arose, examples from the manual were used to clarify. These consensus scores were used as data in the final study. If a tape was double-coded, the scores differed, and it was not then coded by consensus, the author flipped a coin to determine which coder’s rating would be placed in the final database.

Measures

**Measures.** Instruments completed at intake by active client participants for this study included the Experiencing scale, SCL-90-R, Inventory of Interpersonal Problems-CX, Social Skills Inventory, and a demographic questionnaire. The Working Alliance Inventory-Patient Form was completed by the client after sessions one, three, five and seven. At intake the therapists completed the California Psychological Inventory-Empathy Subscale and Sociability Subscale and the Bronx Psychological-Mindedness Scale. Therapists completed the Social Skills Inventory prior to the intake session.
The Experiencing Scale. The concept of experiencing refers to the quality of a person’s participation in therapy. It refers specifically to the extent to which inner referents become the felt data of attention, and the degree to which efforts are made to focus on, expand, and probe those data. Experiencing is a technique for measuring self-involvement in psychotherapy. Experiencing refers to the quality of an individual’s experiencing of himself or herself, the extent to which her ongoing, bodily, felt flow of experiencing is the basic datum of his or her awareness and communications about himself or herself, and the extent to which this inner datum is integral to action and thought. At low levels of experiencing, discourse is markedly impersonal and superficial. Moving up the scale, there is a progression from simple, limited, or externalized self-references to inwardly elaborated descriptions of feelings. At higher experiencing levels, feelings are explored and higher and emergent levels of experiencing serve as the basic referents for problem-resolution and self-understanding.

The scale is independent from specific pathology or problem content. It is also not dependent on the therapists’ orientation or technique. Experiencing should be relevant for all therapies where change in the patient’s level of expressiveness, self-awareness, or self understanding is a goal, or where self-attitudes and expressive behavior are in any way manipulated as part of the treatment procedure (Klein et al., 1969).

Gendlin, one of the creators of the Experiencing scale, holds that good therapy of any persuasion must be experientially centered. It is important to note that Gendlin is not simply talking about the branch of therapy referred to as experiential. He is talking about experiencing in slightly broader terms that cut across orientations. Most current therapies
employ experiential focusing like that described in the experiencing scale; i.e., free association in psychoanalysis, desensitization in behavior therapy, examination of irrational thoughts in CBT. According to Gendlin, these methods work when they involve experiential content, and fail when they become remote, stale, overly intellectual, or highly structured. Whatever else the client does, his or her attention must be his or her immediate, concrete, fresh, bodily sense or feel of the issues. The approach also defines the therapists’ work as independent of the client’s diagnosis, the client’s problem content, the therapist’s theoretical orientation or technique, features of their interaction, or the emotional climate of the relationship (Gendlin, 1967; Gendlin, 1964). The scale has been applied to therapies of different orientations as well as to formats beyond therapy. It has been applied to monologues, structured interviews, group therapy, and Gestalt two-chair exercises, and for written materials such as personal documents and responses to open-ended questions (Klein et al., 1986). This makes it ideal for a study on FIS, which should be present in any human interaction.

The scale attempts to assess the degree to which the patient communicates her personal, phenomenological perspective and employs it productively in the therapy session. The therapist's task is to help the client expand his or her experiencing by supporting the process of focusing and by responding to his or her implicit referent. An appropriate intervention moves just beyond the client’s current experiential level, directing the client towards the implicit meaning or aspect of her experiencing, so that the patient can embrace it and refer to it more directly. The most accurate or theoretically elegant interpretation may make no impact on the client whatsoever if it is more complex
experientially than the patient’s current approach to the specific concern or if it is too intellectual or too remote from the ongoing feelings. While clinical experience and theoretical knowledge may be invaluable in helping the therapist evaluate the client’s problems and enable the therapist to be aware of the important treatment issues, it is the therapist’s sensitivity to the client’s referent for his or her expressed mode of experiencing that enables the therapist to help the client find the next-most-important thing in his or her experiencing, and thus too communicate with and effectively influence the client (Klein et al., 1969).

Rater reliability for psychotherapy sessions has been consistently high. Despite variations in therapist orientation and client problem, correlation coefficients were in the 80’s or 90’s for 12 of 15 published studies reviewed by Klein. The range of these studies was .65 to .91 for mode and .61 to .93 for peak scores. The mean of these 15 studies was .83 for mode and .85 for peak. No variations due to segment length or data form (tape vs. transcript) were apparent (Klein et al., 1986). The validity of the scale has been supported by its correlation with some client variables, such as introspectiveness and cognitive complexity (Klein et al., 1986).

The Experiencing scale consists of one 7-point scale designed to be applied to tape recordings or transcripts of psychotherapy. The seven scale “stages” or steps define the progression of client involvement in inner referents from impersonal (1) or superficial (2), through externalized or limited references to feelings (3), to direct inner referents (4), to questioning an unclear inner referent (5), to focusing with a step of resolution (6), and
finally to the point where focusing comes early and provides the connections for inner discourse (7) (Klein et al., 1986).

The coder assigns a mode and peak score for each segment. The modal rating characterizes the overall, general or average scale level of the segment or unit. It is representative of the general experiencing level in the segment. The peak rating is given to the highest level reached in the segment. Often a peak is found in only a brief segment of the material. In some segments there may be no clear peak and the modal and peak ratings are identical (Klein et al., 1986).

In order to capture the experiencing level across sessions, several variables were created. The mean-mode is the mean of the modes for the 4 sessions, and similarly the mean-peak score is the mean of the peaks for the 4 sessions. The peak-mode score is the highest mode score which was achieved in the 4 sessions. The peak-peak score is the highest peak score in all sessions.

**Demographic questionnaire.** This 8-item questionnaire asked patients to list their age, gender, college major, year in school, race, psychotherapy treatment history (i.e., # of sessions of therapy attended in the past) and the income level of their parents.

**Outcome Questionnaire (OQ).** Prior to intake, all subjects in the study will complete the OQ (Umphress, Lambert, Smart, Barlow & Clouse, 1977) as a measure of pre-treatment psychological distress and will complete it again after sessions two and three as an index of symptom change. The OQ was designed as a measure to track progress and outcome on a session by session basis. It is a 45-item self-report measure that requires respondents to rate their feelings on a broad range of items indicative of general well-
being and psychological functioning. Internal consistency of the OQ is high (coefficient alpha = .93) and there is support for its construct validity. Within a university counseling center, the OQ was shown to have concurrent validity of .78 with the SCL-90-R.

Residualized OQ change scores for both termination and follow-up time points were calculated. The termination OQ change score measured residualized change from the intake OQ to the termination OQ, and the follow-up OQ change scores measured change from the intake OQ to the follow-up OQ at 20 months.

**Inventory of Interpersonal Problems (IIP).** The IIP (Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988) is a 127-item self-report measure that details different types of interpersonal problems people often report and the level of distress accompanying each problem. The items fall into two categories, behaviors that are hard to do (78 items) and behaviors that are done too much (49 items), corresponding to typical patient reporting patterns noted in psychotherapy. The degree of distress associated with each item is rated on a 5-point Likert scale. Items fall into one of six empirically derived subscales including: Hard to be Assertive, Hard to be Sociable, Hard to be Intimate, Hard to be Submissive, Too Responsible, and Too Controlling (Horowitz, et. al 1988).

The present study will employ the 64 item IIP-CX (Alden, Wiggins, & Pincus, 1990), a short form version of the original scale which contains items which match interpersonal circumplex scales. Factor analysis of the IIP-CX revealed subscales (Domineering, Vindictive, Overly-Cold, Socially Avoidant, Nonassertive, Exploitable, Overly-Nurturant, and Intrusive) of eight items each which fall into different octants within Wiggins’ two dimensional circumplex model (1982). For the purpose of analyses,
patients’ scores were collapsed into four of these subscales/quadrants: domineering, overly-nurturant, nonassertive, and overly-cold.

The IIP has adequate convergent and discriminant validity with other self-report, interpersonally focused measures including the SCL-90-R (Derogatis, 1983)(r =.64, correlation of total scores), UCLA Loneliness Scale (Scale L; Russell, Peplau, & Cutrona, 1980) (r’s range from .37 to .73 for separate IIP subscales), Rathus Assertiveness schedule (Scale A; Rathus, 1973) (r’s range from -.16 to .64 for separate IIP subscales), and the Interpersonal Dependency Inventory (Scale D; Hirschfield, Klerman, Gough, Barrett, Korchin, & Chodoff, 1977) (r’s range from .28 to .58 for separate IIP subscales) (Horowitz et. al, 1988).

Residualized IIP change scores for both termination and follow-up time points were calculated. The termination IIP change score measured residualized change from the intake IIP to the termination IIP, and the follow-up IIP change scores measured change from the intake IIP to the follow-up IIP at 20 months.

**Social Skills Inventory (SSI).** The SSI (Riggio, 1989) is a 90-item self report instrument which assesses basic verbal and nonverbal social communication skills. Social communication skills fall into six empirically-derived domains or subscales: Emotional Expressivity (EE), Emotional Sensitivity (ES), Emotional Control (EC), Social Expressivity (SE), Social Sensitivity (SS), and Social Control (SC). These yield a total score which reflects a global level of skill development, social competence or social intelligence. The total score measures were used for the purpose of analyses of patient-completed SSI’s. A mean split was used in ANOVAs as therapist SSI total scores are
range restricted. Therapists were chosen based on the extremity of their social skills (e.g., very high or low).

Measures of test-retest reliability for the SSI range from .81 to .96 for a two-week interval, and alpha coefficients for the SSI subscales range from .62 to .87 (Riggio, 1989). Convergent and discriminant validity for the SSI were supported in a series of studies conducted by Riggio (1986). Subscales EE, ES, SE, and SC were significantly positively correlated with the Affective Communication Test (ACT; Friedman, Prince, Riggio, & DiMatteo, 1980) (r =.59, .44, .67, and .68, p <.001), respectively. In addition, subscales ES and EE were significantly positively correlated with the Profile of Nonverbal Sensitivity (PONS; Rosenthal, Hall, Dimatteo, Rogers, & Archer, 1979) (r =.19 and .18, p <.05), respectively.

**SCL-90-R.** The SCL-90-R (Derogatis, 1983) is a 90-item psychological symptom self-report measure. Items represent problems that have distressed patients during the past 7 days. Each item is rated on a 5-point scale of distress with a 0 response indicating “not-at-all” and a 4 response indicating “extremely”. Items comprise 9 primary symptom dimensions: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, psychoticism, and three global indices of distress: Global Severity Index (GSI), positive symptom distress index, and positive symptom total. The GSI, a cutoff measure used in the present study, provides a sensitive numeric indicator of the participant’s overall psychological distress, integrating information regarding the quantity and intensity of symptoms.
Reliability measures for the 9 primary symptom dimensions consist of internal consistency coefficient alphas ranging from .77 to .90 and test-retest reliability coefficients ranging from .78 to .90 for a one week interval. The SCL-90-R has been shown to reflect a high degree of convergent validity in studies comparing its symptoms scales to the MMPI clinical and content scales. Correlation coefficients ranged from .42 to .64 for the MMPI clinical scales and from .40 to .75 for the MMPI content scales (Derogatis, 1983).

Working Alliance Inventory (WAI). The WAI (Horvath & Greenberg, 1986) is a 36-item self-report measure based on Bordin’s (1979) transtheoretical formulation of the therapeutic alliance. The WAI is comprised of three empirically derived subscales: Task Agreement, Goal Agreement, and Bond Development which each yield individual subscale scores in addition to one combined overall index score. Each subscale contains twelve items scored on a seven-point Likert scale.

High internal consistencies have been reported for the complete WAI (e.g., coefficient alpha of .96 for WAI-C and .95 for the WAI-T; Tichenor & Hill, 1989) as well as individual subscales (e.g, alpha coefficients range from .85 to .88 for subscales of the WAI-C and from .68 to .87 for subscales of the WAI-T; Horvath & Greenberg, 1989). In addition, multiple findings support the convergent and discriminant validity of the WAI Goal and Task scales, but evidence regarding the convergent validity of the Bond scale is equivocal (Greenberg & Pinsof, 1986; Horvath & Greenberg, 1989). However, some studies have failed to report significant correlations between WAI-T and WAI-C scores, suggesting that measurement of the alliance from different perspectives can yield
different results (Tichenor & Hill, 1989). Horvath and Greenberg (1989) reported the WAI-C to demonstrate strong predictive validity given the scale’s significant correlations with counseling outcome measures such as the Patient Post-therapy Questionnaire (CPQ; Strupp, Wallach, & Wogan, 1964), State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970) and the Scale of Indecision (SI; Osipow, Carney, & Barak, 1976).

Mean client and therapist WAI scores over therapy were computed by averaging client and therapist WAI scores for sessions 1, 3, 5, and 7.

**Narrative Process Codes (NPC).** The narrative is a two-step process that enables the researcher to: a) reliably subdivide and characterize therapy session transcripts into topic segments according to content shifts in verbal dialogue, and b) further subdivide and characterize these topic segments in terms of one of three narrative-process mode types – External, Internal, Reflexive - that are termed narrative-process sequences for coding purposes:

i. External narrative process sequences that include description of events (past, present, and/or future; actual or imagined)

ii. Internal narrative process sequences that include a subjective/experiential description of experience.

iii. Reflexive narrative process sequences that include analysis of current, past and/or future events (Angus, Levitt, & Hardtke, 1999).

In order to match the narrative process codes which had been completed for session 3 with experiencing coded segments, coders used the session transcripts with line numbers
and the Narrative Process Coding Scoring sheets. They would find the beginning of the experiencing coded segment on the transcript, then find the line number on the page. These line numbers would also be marked on the NPCS scoring sheet. The coder would then calculate the percentage of lines for the experiencing coded segment which were at each of the three levels of the NPCS. The new variables were % External, % Reflexive, and % Internal.

RESULTS

Design

Hypotheses 1-3 used regression to see if alliance, narrative process codes or outcome would predict client experiencing scores. For the hypotheses about FIS and client social skill, a MANOVA was used because the correlations between the four measures of experiencing (mean-mode, peak-mode, mean-peak, peak-peak) were strong for the two mode measures (mean-mode and peak-mode, .857) and the two peak measures (mean-peak and peak-peak, .755). Thus, first the modes, mean-mode and peak-mode, were entered as dependent variables in a MANOVA with the three categorical fixed factors of FIS, client social skill, and therapist training status. Then the peaks, mean-peak and peak-peak, were entered into a MANOVA with the same fixed factors.

Summary

Regression revealed that experiencing scores were not predicted by alliance or outcome, but were predicted by internal and reflexive narrative process codes (See Table 7 for correlations between all variables).
Table 7

Intercorrelations between experiencing and hypothesized variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mean Peak</td>
<td></td>
<td>.884</td>
<td>-.107</td>
<td>.392</td>
<td>.059</td>
<td>.138</td>
<td>.149</td>
<td>.395</td>
</tr>
<tr>
<td>2. Mean Mode</td>
<td></td>
<td></td>
<td>-.119</td>
<td>.344</td>
<td>.180</td>
<td>.062</td>
<td>.096</td>
<td>.443</td>
</tr>
<tr>
<td>3. WAI</td>
<td></td>
<td></td>
<td></td>
<td>.30</td>
<td>.149</td>
<td>.355</td>
<td>.160</td>
<td>-.194</td>
</tr>
<tr>
<td>4. % Internal NPC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.123</td>
<td>.042</td>
<td>.434</td>
<td>.227</td>
</tr>
<tr>
<td>5. IIP Change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.285</td>
<td>.137</td>
<td>-.019</td>
</tr>
<tr>
<td>6. FIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>7. Client Skill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>na</td>
</tr>
<tr>
<td>8. Therapist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--</td>
</tr>
</tbody>
</table>

Note. WAI is average client-rated WAI over 7 sessions. IIP change score was computed by subtracting intake IIP from follow-up. Client skill is client social skill.

The MANOVA’s revealed that trained therapists had clients with higher peak-mode, mean-peak, and peak-peak scores.
Hypotheses I

Hypotheses I, which stated that dyads with higher mean client-rated alliance scores would have higher mean-mode and mean-peak experiencing scores over therapy, was not supported. Mean client-rated WAI scores did not predict the clients’ mean-mode experiencing score $F (1, 44) = .20, p = .66$ or the clients’ mean-peak experiencing scores $F (1, 44) = .01, p = .94$. Post-hoc analyses of the relationship between individual session and average WAI scores for both therapists and clients and all possible experiencing scores, both for individual sessions and across therapy, did not reveal any significant relationships with the alliance.

Hypotheses II

There was some support for hypothesis II which stated that clients who communicate narratives that are more emotionally internal and meaningful will have higher mean-mode and mean-peak experiencing scores, and that clients who communicate narratives that are more external to the therapeutic process will have lower experiencing scores. To test this hypotheses the percentage of time spent in each of the three modes (% External, % Reflexive, % Internal) were entered into a multiple regression model. % Internal and % Reflexive were able to predict mean-peak experiencing, $F (2, 32) = 3.41, p = .046$. This finding provides some support for the validity of experiencing in this study. % External was excluded from the model both times. The $n$ was less than 43 because not all client had session three coded for Narrative Process Codes.
Hypotheses III

Hypothesis III, which stated that clients who had better outcomes will display higher peak-peak levels of experiencing, was not supported. Residualized OQ change scores at follow-up did not predict experiencing scores \( F(1, 35) = 1.27, p = .27 \). IIP follow-up change scores did not predict the peak-peak experiencing scores, \( F(1, 35) = .27, p = .61 \). The number of participants was lower here because nine clients did not fill out the follow-up outcome measures. Post-hoc analyses also examined the relationships between both OQ and IIP termination change scores, OQ and IIP follow-up scores, and all possible combinations of experiencing scores, both for individual sessions and across therapy, and found no significant relationships.

Hypotheses IV & V

For the remaining hypotheses the categorical variables of therapist FIS, client social skill, and therapist training status were entered as fixed factors in MANOVAs. Two MANOVAs were run. First the mode scores, mean-mode and peak-mode, and then the peak scores, mean-peak and peak-peak, were entered as dependent variables. Because the interaction terms were not significant, the analysis was run again and the interactions were suppressed (See table 8 for mean experiencing scores by FIS, client social skill, and therapist training).
Table 8

Experiencing scores and standard deviations for FIS, training status and client social skill

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Mean-Peak</th>
<th>Mean-Mode</th>
<th>Peak-Peak</th>
<th>Peak-Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>High FIS</td>
<td>20</td>
<td>3.14</td>
<td>.55</td>
<td>2.31</td>
<td>.54</td>
</tr>
<tr>
<td>Low FIS</td>
<td>25</td>
<td>2.96</td>
<td>.72</td>
<td>2.24</td>
<td>.64</td>
</tr>
<tr>
<td>Trained</td>
<td>21</td>
<td>3.31</td>
<td>.60</td>
<td>2.55</td>
<td>.48</td>
</tr>
<tr>
<td>Untrained</td>
<td>24</td>
<td>2.80</td>
<td>.61</td>
<td>2.03</td>
<td>.57</td>
</tr>
<tr>
<td>High Social Skill Clients</td>
<td>22</td>
<td>3.14</td>
<td>.72</td>
<td>2.33</td>
<td>.70</td>
</tr>
<tr>
<td>Low Social Skill Clients</td>
<td>23</td>
<td>2.95</td>
<td>.57</td>
<td>2.22</td>
<td>.47</td>
</tr>
</tbody>
</table>
Hypotheses IV, which stated that therapists with high FIS would have clients with higher mode and peak experiencing scores, was not supported by the data. Entering the modes in the MANOVA revealed that high FIS therapists did not have clients with higher mean-modes, F (1, 41) = .09, p = .77 or peak-modes, F (1, 41) = .24, p = .63. Entering the peaks showed that high FIS therapists did not have clients with higher mean-peak, F (1, 41) = .77, p = .39 or peak-peak scores, F (1, 41) = .23, p = .13. (See table 9 for the MANOVA for the modes and table 10 for the MANOVA for the peaks).

Table 9

MANOVA for modes

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>df</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIS</td>
<td>Peak-Mode</td>
<td>1</td>
<td>.24</td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>Mean-Mode</td>
<td>1</td>
<td>.09</td>
<td>.77</td>
</tr>
<tr>
<td>Client Social Skill</td>
<td>Peak-Mode</td>
<td>1</td>
<td>1.4</td>
<td>.24</td>
</tr>
<tr>
<td></td>
<td>Mean-Mode</td>
<td>1</td>
<td>.26</td>
<td>.62</td>
</tr>
<tr>
<td>Training</td>
<td>Peak-Mode</td>
<td>1</td>
<td>3.77</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>Mean-Mode</td>
<td>1</td>
<td>9.68**</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**p < .01.
Table 10

MANOVA for peaks

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>df</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIS</td>
<td>Peak-peak</td>
<td>1</td>
<td>.77</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>Mean-peak</td>
<td>1</td>
<td>2.34</td>
<td>.38</td>
</tr>
<tr>
<td>Client Social Skill</td>
<td>Peak-peak</td>
<td>1</td>
<td>.08</td>
<td>.78</td>
</tr>
<tr>
<td></td>
<td>Mean-peak</td>
<td>1</td>
<td>.88</td>
<td>.35</td>
</tr>
<tr>
<td>Training</td>
<td>Peak-peak</td>
<td>1</td>
<td>5.2*</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>Mean-peak</td>
<td>1</td>
<td>7.17*</td>
<td>.01</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05.

Hypothesis V

Hypotheses V, which stated that high social skill clients would have higher mean and peak levels of experiencing than low social skill clients, was not supported. In the MANOVA with the modes entered as dependent variables, high social skill clients did not have higher mean-modes, F (1, 41) = .26, p = .62 or peak-modes, F (1, 41) = 1.40, p = .24. In the MANOVA for peak scores, higher social skill clients did not have higher mean-peaks, F (1, 41) = .88, p = .35 or peak-peak scores, F (1, 41) = .08, p = .78.
Additional Analyses

The MANOVA testing modes as dependent variables revealed that trained therapists did not have clients with higher mean-modes, $F (1, 41) = 3.77, p = .06$, than untrained therapists. However, trained therapists did have clients with higher peak-modes, $F (1, 41) = 9.68, p = .00$. The MANOVA for peaks showed that trained therapists had clients with higher mean-peak, $F (1, 41) = 7.17, p = .01$, and peak-peak scores, $F (1, 41) = 5.2, p = .03$.

DISCUSSION

Summary of Results

In this study four of five hypotheses were not supported. Experiencing failed to show significant relationships to the client-rated WAI or outcome as measured by the IIP and OQ. High and low social skilled clients did not differ on their levels of experiencing. High and low FIS therapists also did not differ on the level of experiencing among their clients.

Experiencing was predicted by internal and reflexive Narrative Process Codes. Finally, trained and untrained therapists did differ on the level of experiencing shown by their clients. Possible explanations for these results follow.

Exploring Null Findings

Because some of the non-significant findings were in the predicted direction, we conducted a power analysis to see more subjects would have helped. The program Sample Power program indicated that an extremely large n, especially for a psychotherapy research study, would have been necessary to detect findings for most of
the hypotheses. Entering the most significant alliance and experiencing relationship, an n of 779 participants would have been required to provide power of 80%. Given the most significant outcome-experiencing relationship, which was between peak-peak and OQ termination, a n of 191 would have been needed to provide power of 80%. An n of 64 would have been required to achieve 80% power to detect the FIS effect in the MANOVA. A n of 800 would have been required to achieve 80% power with client social skill. Thus, it seems that only FIS had a feasible n of 64 to achieve 80% power.

Another problem may have been a low range of scores on the experiencing scale. Klein et. al (1986) state that a low range or a generally low level of experiencing ratings may contribute to a lack of effects. This problem of a restricted range has been found among the group at York University who have used the experiencing scale in their studies of emotion focused therapy. They typically find experiencing codes in the range of 2-3.5 (Greenberg, 2001). A study by Rubenstein (1970) divided therapists into high and low facilitators based on their ratings on the Accurate Empathy, Non-possessive Warmth, and Congruence scales. Rubenstein found no experiencing differences for facilitation level and reasoned that it might have been due to the low range of experiencing ratings (average 2.63, range from 1.00 to 4.80).

A restricted range of scores can attenuate correlations. In this thesis the range was only slightly greater than in Rubenstein’s study. The range for modes was from 1-5 with a mean of 2.30 and the range for peaks was 1-6 with a mean of 3.05. Only one segment received a peak of six. It is possible that this restricted range contributed to a lack of findings as it may have in the Rubenstein study. Figure 1 shows the frequencies of experiencing ratings in the final data set.
Hypotheses I—Working Alliance Inventory

Hypotheses I, which predicted a positive relationship between mean client-rated WAI scores and mean-mode and mean-peak experiencing scores over the course of therapy, was not supported. This hypothesis was based on findings that alliance and experiencing have both been found to relate to outcome (Castonguay et al., 1996). However, it is possible that the relationship between them is complex and also depends on the moment to moment needs of the client. Below several studies will be presented which hint that at different times, the goals of building and maintaining an alliance and facilitating experiencing may not be perfectly aligned.

Kivlighan and Schmitz (1992) studied counselor technical activity in dyads in which alliances were categorized as either improving or continuing as poor based on client-rated
WAI scores. In the improving dyads the counselors, who were all trainees, were relatively more challenging versus supportive than in the continuing poor alliance dyads. Challenging is not an intervention style found in the literature describing methods for increasing experiencing (Gendlin, 1996). Kivlighan and Schmitz go on to state that counselors who use too much support by smoothing over rough spots in the relationship may be further complicating the alliances. Addressing these rough spots may strengthen the long-term alliance, but it might take the focus away from facilitating experiencing.

In an earlier study by Kivlighan et. al (1990), several interventions which have been associated with experiencing showed a negative relationship with the client-rated WAI. This study studied the effect of counselor intentions (Hill et al., 1988) on client-rated WAI. The interventions studied were set limits, assessment, explore, support, restructure, change, and educate. Surprisingly support was negatively related to the alliance. Intuitively it seems that support would be related to the alliance, and it is also theorized to aid experiencing (Gendlin, 1996). Kivlighan et. al theorized that support may place the client in a passive role, and that the best alliances occur when the client takes a more active responsibility for the content and direction of the interaction.

In the same Kivlighan et. al study the explore intention, which has a focus on feelings, thoughts and behaviors was also negatively related to WAI scores. Placing the focus on feelings is largely what experiencing is about, and therefore this shows that experiencing related constructs are also not always related to the alliance. The authors hypothesized that the use of this intention early in therapy may have prematurely focused the client on areas with which he or she was not ready to deal, thus detracting from the formation of
the alliance. Again the therapists may need to decide when to be alliance focused and when to be experiencing focused.

Another example of the interactions between counselor technical activity and the alliance is a study by Horowitz et. al (1984) which found that when clients were highly motivated, the alliance was strengthened by an examination of negative transference. However, when clients were poorly motivated, the alliance was strengthened by a consistently positive attitude on the part of the therapist. This is an example of the research area designed to match therapists to clients and match therapist interventions to their clients needs. The theory is that different clients require different therapies at different times.

William Stiles (1994) argued that the requirements and responsiveness of clients helps explain some of the difficulties of producing findings in psychotherapy research. Stiles explained that with psychotherapy interventions, the more-is-better, drug metaphor is not applicable. This metaphor implies that if an intervention is good for the client, much like aspirin is good for pain, then more of the intervention will deliver a better outcome for the client, yielding a positive correlation between the intervention and outcome. Stiles argues that this metaphor does not apply in psychotherapy because it overlooks therapist responsiveness to varying client requirements for process components. Appropriate responsiveness tends to defeat the process-outcome correlation. A client may be doing great and thus not require much of an effective intervention. In this case there would be a negative correlation between outcome and the intervention, even if the intervention is theoretically helpful. In a similar vein, a therapist may provide many interventions if a client starts deteriorating. It is possible that an effective intervention would not be enough
to pull the client out of a poor outcome. In this case there would again be a negative relationship with the intervention and outcome, as the therapist offered many of them to the client who was headed towards a poor outcome. This may help explain why experiencing and the alliance are not related in this study. The complicated needs of clients and the agendas of the therapists may have led them to address the alliance and experiencing at different times and in differing amounts depending on the needs of the client.

**Hypotheses III- Outcome**

Hypotheses 3, which stated that there would be a positive relationship between outcome as measured by the IIP and OQ at follow-up and peak-peak experiencing scores, was not significant. Post-hoc analysis, which aimed to determine if there were relationships between either mode or peak experiencing scores of any one session and outcome, revealed a few trends. The relationship between the termination IIP change score and the session 2 mode score was almost significant, $F (1, 44) = 3.76, p = .059$. The relationship between the OQ follow-up change score and the session 3 peak score was also close to significant, $F (1, 35) = 3.69, p = .063$. There seems to be slight indicators of a trend between early session experiencing and the outcome variables, but not a very strong one.

Some previous research might suggest that the relationship between outcome and experiencing is not always that clear, possibly explaining the overall lack of findings. Bommert and Dahlhoff (1978) did not find a significant difference between second session experiencing levels of more and less successful clients. Luborsky (1982) failed to find any overall relationships between experiencing in the first two sessions of
psychodynamic therapy with either raw, rated or residual improvement scores. Richert (1976) found ratings of experiencing in the eight middle sessions to be unrelated to cognitive change measures and to be negatively related to self-satisfaction at the end of therapy. Custers (1973) found no association for experiencing and either MMPI or Butler-Haigh Q-sort outcomes with two segments sampled from the middle of therapy.

In general psychotherapy process-outcome research findings have often been inconsistent and disappointing (Elliot, 1986; Greenberg & Pinsof, 1986; Orlinsky & Howard, 1986; Shapiro et al., 1994). The one exception to this area seems to be the consistent finding linking the alliance to outcome, as was found in the OUHRS. Other process linkages are often elusive. Although some studies report significant relationships, these are seldom reliably replicated (Stiles & Shapiro, 1994). According to Stiles, the best designed studies have found the weakest relationships. For example, Moras, Elkin, Imber, Watkins, and Sotsky (1989, May) developed measures specifically for the theory of interpersonal therapy as was practiced in the NIMH TDCRP. They measured both the delivery of theoretically crucial components and uptake by clients and compared the processes of good and poor outcome clients. After careful sampling and measurement they found negligible statistical relationships between process and outcome.

Once again, Stiles and Shapiro’s (1994) refutation of the drug metaphor may elucidate some of the difficulties in finding process-outcome findings. The process-outcome strategy and the drug metaphor overlook therapists’ and clients’ responsiveness to each other and the evolving therapy interaction. Process variations are likely to reflect moment by moment adjustments to participants’ requirements, and should not be considered as either planned or random manipulations of treatment strength (Stiles & Shapiro, 1995).
Applying this logic to the OUHRS, a certain client may have required more experiencing because his or her condition was declining, and the therapist may have been able to facilitate more experiencing in this client. In this situation experiencing would have a negative relationship to outcome, even though it may have been a beneficial intervention for the client. These explanations of the overall difficulty of process-outcome research provide some insight into how this study.

Hypothesis IV - Therapist FIS

There was no support for Hypotheses IV, which stated that high FIS therapists will have clients with higher mean-mode, peak-mode, mean-peak and peak-peak experiencing scores than low FIS therapists. Table 8 shows that three of the four scores for high and low FIS therapists were in the predicted direction, with high FIS therapists having clients with higher experiencing scores, but the results were not significant.

Some studies have also failed to show large relationships between variables such as empathy, which is included in the FIS construct, and experiencing. Jennen, Lietaer, and Rombauts (1978) studied 176 segments from the beginning, middle and end of therapy for 13 patients and found no overall correlation of rated accurate empathy and experiencing. Within cases, accurate empathy and experiencing were significantly correlated for only 4 of the 13 patients.

Rubenstein (1970) had 16 experienced therapists submit tapes of their work with three to four clients each, from which he selected nine 4-minute segments to make ratings of accurate empathy, nonpossessive warmth, and congruence ratings. On the basis of these findings, therapists were then classified as either high or low facilitators. Rubenstein then located examples of eight different categories of therapist behavior and extracted client
speech segments (which varied in length from \( \frac{1}{2} \) to 2 minutes) from points before and after each type of therapist intervention. Experiencing ratings of these segments showed no significant main effects or interactions for therapist facilitation level. The only interaction that approached significance \((p < .08)\) suggested that high facilitative therapists responded at points where client experiencing was higher more often than did low facilitators, but that these responses lowered client experiencing.

There have also been studies showing that experiencing is best facilitated with a focused and consistent approach. Thus, high skill alone may not be enough if the therapist is not consistently focused on facilitating experiencing. In reviewing the literature Klein (1986) stated that experiencing is best facilitated when the therapist is directly focused on the experiential process. A study by Olsen (1975) revealed how a focused intervention can raise experiencing, but it also revealed how experiencing can drop off when the focus on experiencing is lost. Olsen administered an adaptation of Gendlin’s (1986) focusing procedure that included visual imagery and relaxation to 23 nonpsychotic outpatients who were typical of private practice outpatients. Training in Gendlin’s focusing techniques, designed to facilitate experiencing, was provided to therapists. Experiencing significantly improved over pretraining therapy sessions, and most clients reached stage 4 or above with focusing instructions. The ability to focus and reach these higher stages of experiencing was also significantly correlated with the clients’ reports of mastery and insight on the Therapy Session Report. However, when focusing instructions were not given in regular therapy sessions after training, changes in experiencing were not maintained. This study provides more evidence that therapists might need to direct their attention to experiencing to keep it elevated.
Jachim (1978) attempted to raise experiencing with a role induction interview or filmed modeling in low SES community mental health center patients. Neither intervention had any effect on experiencing, self-disclosure, or attraction to therapy. The authors felt that this may have occurred because the induction touched upon many issues, possibly limiting the potential focus on experiencing. It appears that attempts to facilitate experiencing work best when that is the only goal of the session/intervention.

The OUHRS therapists had no specific instructions except to be helpful to their clients, and it is possible that the high FIS therapists had several objectives besides raising experiencing. Half of the high FIS therapists were untrained and were not ingrained in a culture that states that raising experiencing is beneficial to therapy outcome, and thus should be a focus of the therapy hour. This finding is explored in a later section. As a whole the high FIS therapists found a way to increase outcomes more than the low FIS therapists without focusing on experiencing.

**Hypothesis V - Client skill**

This hypothesis, which predicted that clients with high social skills will have higher mean-mode, peak-mode, mean-peak and peak-peak levels of experiencing than low social skill clients, was not supported. In a post-hoc analysis the strongest relationships between either mode or peak experiencing scores for any one session and client social skills appeared for session. The relationship between client social skills and session 4 mode approached significance, F (1, 44) = 3.31, p = .076, as did the relationship between client social skill and session 4 peak, F (1, 44) = 3.41, p = .072. Thus, again we have some indications of a possible trend, but it required looking at all possible combinations of
experiencing scores and sessions to find it. No statistical adjustment was made for test-
wise error, as the tests were not significant.

The literature gives us clues as to how this could occur. The hypotheses that high
client social skill would lead to higher experiencing was based on early research which
suggested that experiencing was to some degree a measure of good mental health as well
as an index of productive therapeutic involvement (Kiesler, 1971; Kiesler et al., 1965;
Rogers et al., 1967). Experiencing was also related to various indicators of verbal and
expressive capacity that have been consistently associated with good motivation and
prognosis for therapy (Rogers et al., 1967). Given that the high social skill clients in the
OUHRS had better outcomes, it seemed reasonable to predict that they would have higher
experiencing.

However, some other studies may explain why client social skill was not strongly
related to experiencing. The relationship of experiencing to measures of health,
personality, and cognitive style have not been consistent. In one study experiencing was
more uniformly associated with neuroticism as a general factor than it was with any
specific neurotic subtype or category of affective distress (Fontana, Dowds, & Eisenstadt,
1980). There was some additional evidence of experiencing having been higher in
concert with introspectiveness (Gendlin, Beebe, Cassens, Klein, & Oberlander, 1968),
obsessiveness, or self-consciousness in both help seeking and non help-seeking samples
(Kiesler, 1969). Finally, experiencing has also been associated with measures of
cognitive style, such as complexity and differentiation, as well as with other indicators of
reflectiveness, expressive capacity, or attraction to psychotherapy (Klein et al., 1986).
Thus, we see that the relationship between client variables and experiencing has often
been confusing and equivocal, possibly explaining the lack of a relationship between social skill and experiencing.

*Exploring the effects of Training on Experiencing*  

To further understand these results, possible explanations for how training facilitated higher experiencing were explored. Psychology’s tradition of valuing experiencing, the relationship between various interventions and experiencing, and the role of training programs on therapist intervention patterns and behaviors that facilitate experiencing were examined. After exploring the link between training and experiencing, the lack of a link between experiencing and outcome in this study were investigated. The literature on experiencing was reviewed, including studies which hint that experiencing may be more tied to specific interventions or orientations. A case will be made that experiencing is not always a common factor, and that it may be more tied to specific orientations and interventions. As the Dodo Bird verdict has shown, specific orientations and interventions do not always reveal clear links to outcome, possibly explaining why experiencing was not predicted by outcome. Finally, possible remaining contributors to outcome will be explored. The case for the alliance and FIS may be even stronger, now that another process has failed to show its contribution to outcome. Possible methods for further exploration of these two constructs will be discussed.

How or why did training facilitate experiencing?

First we must examine how therapist training effected ratings of client experiencing. Was there something in the training of these therapists that enabled and/or encouraged them to facilitate experiencing? In essence this is asking two questions: does therapist
training teach therapists to value and pursue deeper experiencing, and does it develop skills or techniques to achieve this aim?

Did the trained therapists have the desire to facilitate experiencing?

In a post-hoc analysis, the luxury of asking the therapists what their intentions were is not available. The original study also did not collect therapists opinions on what they perceived to be meaningful moments in therapy. To attempt to gain a glimpse of the mindset of the trained therapists, we can explore some of the more widely held beliefs in the field, as well as aspects of their specific training program.

As described in the introduction of this thesis, many common factors models have posited that experiencing or similar processes are beneficial in therapy. Frank and Frank (1991) state that the client’s emotions are aroused in successful therapy. Castonguay (1993) writes that successful therapy depends on common factors such as insight, corrective experiences, the opportunity to express emotions and the acquisition of a sense of mastery, all of which would seem to correspond with high experiencing. Grencavage & Norcross’s (1990) model includes the opportunity for catharsis and ventilation and a focus on one’s "inner world"-with exploration of emotional issues. Lambert & Ogles’ (in press) model includes support factors which include therapist/client active participation, therapist warmth/respect/empathy/acceptance/genuineness, catharsis, and learning factors such as insight, corrective emotional experience and affective experiencing.

According to these authors, experiencing or emotional processing is one of the core elements of the therapeutic process regardless of orientation. It is possible that this
belief, that appears to be widely held in the field, may filter down into the psyches of burgeoning therapists.

What influences experiencing?

Are there certain techniques or interventions that facilitate or increase experiencing?

Wiser and Goldfried (1998) coded eighteen sessions of CBT and 13 sessions of psychodynamic-interpersonal therapy obtained from experienced clinicians in a naturalistic setting. They found that reflections and acknowledgments, affiliative and noncontrolling interventions, or interventions highlighting nonspecific client content were associated with maintained high experiencing. Lengthier interventions and interventions rated as affiliative but moderately controlling were associated with shifts to lower experiencing. For clients of CBT therapists, questions, interventions rated affiliative but controlling, and highlighting minimal emotional content were also associated with shifts to lower experiencing. Apparently certain therapist interventions that are more aligned with a given orientation can facilitate or decrease experiencing. So far evidence has been presented that experiencing is valued in various psychotherapy traditions and that certain behaviors increase experiencing. The next question to examine is whether training programs succeed at developing these behaviors.

Does training change therapist behaviors?

While the benefits of training on therapy outcomes is an equivocal research area (Auerbach & Johnson, 1977; Berman & Norton, 1985; Beutler et al., 1994; Christensen & Jacobson, 1994; Durlak, 1979; Hattie, Sharpley, & Rogers, 1984; Nietzel, Russell, Hemmings, & Gretter, 1987; Stein & Lambert, 1995), there is some literature showing
that training does at least influence therapist behaviors. Therapist values, skills and behaviors have been shown to change during the course of formal training (Thompson, 1986) and with clinical experience (Guest & Beutler, 1988; Tracey & Hays, 1989). Studies have shown that students tend to adopt the theoretical orientation of their supervisors (Beutler & McNabb, 1981; Guest & Beutler, 1988).

Even short-term training programs in the form of videos have been shown to alter therapist behaviors. One study evaluated the effectiveness of a teaching package consisting of a booklet and three videotapes to help trainees in psychiatry learn the techniques specific to a conversational model of psychotherapy prior to group supervision. As a result of training there were significant improvements on most of the key skills, and 9 of the 12 trainees improved considerably (Maguire, 1984).

A number of studies have shown that training increases behaviors that have been associated with facilitating experiencing in clients. Hill (1981) found that doctoral students increased their use of minimal encouragers and decreased their use of questions as a result of their training. In simulated interviews untrained “therapists” made directive interventions rather than exploring patients’ feelings or experiences (D'Augelli, Danish, & Brock, 1976). Bohn (1965) found that in responding to tape-recorded enactments of a “hostile,” “dependent,” or “typical” client, relatively experienced graduate student therapists most often used responses categorized as restatement of content or clarification of feelings, whereas naïve undergraduate “counselors” relied most heavily on reassurance, persuasion, direct questioning, and forcing the topic. The use of minimal encouragers, restatements, exploration and clarification of feelings and a tendency away
from directiveness are all often associated with the facilitation of experiencing (Gendlin, 1996).

In process analyses of the Vanderbilt I study, which was one of the forbearers for the OUHRS and this thesis, the professional therapists focused on examination of the patients’ feelings to a greater extent than “inherently helpful” college professors (Gomes-Schwartz & Schwartz, 1978). Similarly, a qualitative analyses of the therapists in the current study revealed that the trained therapists did in fact exhibit different intervention patterns, which are more often associated with higher experiencing, than the untrained therapists (Anderson et. al, 2001).

Evidence that the trained therapists used different techniques

In order to further examine the OUHRS results, a research informed qualitative analysis of therapy sessions was conducted. Six therapists whose outcomes represented the overall findings of the study were chosen for further study. The three therapists with high FIS whose clients improved the most and the three low FIS therapists whose clients improved the least or had negative outcomes were selected to be studied in depth. Each of the three therapists was assigned to two raters, who then watched session 3 on video and listened to all the other sessions for that dyad. The raters took detailed qualitative notes and then compared their findings with each other before presenting them to the group. After all the notes on all therapists were presented, the team attempted to draw some conclusions about the therapists and therapies (Anderson et. al, 2001). Later, based on audience requests at a conference, the team decided to examine the therapies of therapists who did not fit within the main study’s findings. Thus, the team now examined one high
FIS therapist who received a poor outcome and a low FIS therapist whose client did well (Anderson et. al, 2001b).

Although training status was not used to select any of the cases in these qualitative studies, the rating team did make observations about the way in which the trained and untrained therapists conducted therapy. After watching or listening to the entire course of therapy, the team agreed that training status was clearly discernible from the therapist’s style and use of interventions. The trained therapists used more reflections, less self-disclosure and spoke less frequently. The untrained therapists exhibited behaviors which may be deemed unconventional according to various traditions of psychotherapy. One untrained therapist favored one intervention mode, questioning, to the exclusion of almost all others. Untrained therapists used self-disclosure in manners that are often warned against in graduate school. Untrained therapists were more willing to discuss events outside the therapy hour, such as shopping for shoes, which the raters labeled ‘tangents’. Untrained therapists engaged in advice, edification and persuasion more frequently than trained therapists. Overall, the team concluded that the untrained therapists were more likely to engage in what the research team called “buddy therapy.” This involved engaging in a reciprocal conversational style, disclosing, bonding, giving advice and using the common therapist tools of restatements and reflections less often than the trained therapists (Anderson et. al, 2001b). This may help us explain why the trained therapists facilitated higher experiencing.

So far evidence has been presented suggesting that: 1) experiencing is valued by large parts of the field of psychology, 2) therapist behaviors influence experiencing, and 3) the trained therapists were indeed performing different interventions and interventions more
aligned with higher experiencing. Unfortunately for these trained therapists, and possibly for practitioners as a whole, there are no easy formulae which dictate how specific intervention patterns lead to higher outcomes. Indeed, although the trained therapists facilitated higher experiencing, they did not deliver higher outcomes. This leads to an examination of why experiencing did not seem to relate to outcome in this study.

Experiencing may not be so common

Perhaps experiencing is not so common. To be common, a factor should cut across therapies and contribute to outcome regardless of orientation. While it is true that most common factors models include a factor such as emotional processing or insight which could fall under the rubric of experiencing, it is also true that not all orientations pursue experiencing to the same degree. If experiencing is indeed more technique or orientation related, we might gain an understanding of why experiencing did not lead to higher outcomes in this thesis. Many research attempts have been made to link various techniques with outcomes, only to yield inconclusive results (Orlinsky & Howard, 1986; Shapiro et al., 1994; Stiles & Shapiro, 1994).

The experiencing scale emerged from the humanistic tradition, thus it does have links to a specific orientation. Gendlin (1996) has also developed focusing techniques designed to enhance experiencing, with the desire to improve outcomes. He details specific ways for therapists to engender effective client movement, particularly in those difficult times when nothing seems to be happening. While he does always place a strong emphasis on the relationship, both as a constant presence and one that varies throughout therapy, he also details more specific therapist responses which can stimulate and enable a client’s
capacity for direct experiencing. Thus, one of the originators of the experiencing concept and scale seems to feel that it is technique related, although it always occurs within and is tied to the relationship.

In one study Jones (1993) used the Process Q-Set to show that although features were common to both CBT and psychodynamic treatments, there were some important differences, particularly related to affect. CBT therapists promoted control of negative affect through the use of intellect and rationality combined with vigorous encouragement, support, and reassurance from therapists. Psychodynamic therapists placed an emphasis on the evocation of affect, on bringing troublesome feelings into awareness, and on integrating current difficulties with previous life experience, using the therapist-patient relationship as a change agent. From these results, it appears that psychodynamic therapists place a higher value on experiencing.

This brings up the issue of how common is experiencing? Experiencing and the other factors which bring it about, such as empathy, may be more important to certain groups (trained vs. untrained) or certain orientations (psychodynamic vs. behavioral). While this one study may not counter the number of studies which have found that experiencing is beneficial to outcome, it may be enough to encourage questioning assumptions about therapy. Wiser and Arnow (2001) write that across orientations experiencing is assumed to be beneficial to outcome. It is this assumption that may need to be challenged.

It may even be possible that in past experiencing studies, experiencing was related to outcomes in studies of trained therapists because all the therapists were trying to facilitate experiencing. Some studies have compared therapists who were all performing the same type of therapy, one that valued and attempted to facilitate experiencing (Bierman &
Lumley, 1973; Elliot, Cline, & Shulman, 1982; Elliot, Klein, & Mathieu-Coughlan, 1983; Hinterkopf & Brunswick, 1981; Jennen et al., 1978; Olsen, 1975; Rogers et al., 1967; Rubenstein, 1970; Schoeninger et al., 1967; Truax & Mitchell, 1971). Because the therapists were all trying to facilitate experiencing, the ones with the better alliances would be the ones able to achieve this goal. Thus, the alliance would correlate with experiencing, which would correlate with outcome. However, it could be the alliance which is driving both. The therapists who were able to meet the sub-goal of increasing experiencing would most likely be able to meet the goal of increasing the alliance and outcome.

Wiser and Arnow (2001) reviewed the literature on experiencing in an attempt to delineate when it may be advisable to facilitate or regulate experiencing. Although it has generally been assumed that experiencing is a beneficial process in therapy, Wiser and Arnow indicate that experiencing is a process that should be regulated and is not always beneficial. Wiser and Arnow believe that it is best to judge on a case by case basis which clients would benefit from deeper experiencing. The three types of clients who warrant extra caution are those who selectively or pervasively avoid experiencing states, those who behave maladaptively, and those whose meaning frameworks have been damaged.

Certain diagnostic groups seem to have these issues at their core. PTSD and bereavement strongly revolve around lost meaning. Individuals with schizoid personality disorder and some with anorexia nervosa appear to avoid experiencing, and individuals with antisocial personality disorder behave maladaptively. Certain clients may become overwhelmed with emotional situations. An extreme example of a client with difficulties regulating emotion would be a borderline client. Linehan’s dialectical behavior therapy,
which has shown some beneficial results with borderline clients, emphasizes containment, management, and reduction of emotional experiencing, not its facilitation (Linehan, 1993). Facilitating experiencing may lead to increased emotional intensity. If this overwhelms or confuses the client, he or she will not be able to benefit from the meaning making potential of experiencing.

Wiser and Arnow also indicate that facilitating experiencing is not advisable with clients who use destructive strategies to cope with heightened experiencing. The ultimate goal is to shift them towards adaptive coping strategies. If therapists heighten experiencing too much before developing client coping skills, the clients may use maladaptive strategies such as abusing substances, binge eating or self-mutilating. Thus, by heightening experiencing, the therapist could be encouraging the acts the client came into therapy to cure. Only when adaptive strategies are made available to, and mastered by, such clients would facilitating experiencing seem indicated.

Wiser and Arnow also agree with many other authors that experiencing should take place in a safe relationship. While experiencing can be exciting and full of learning potential, it can also be unsettling and frightening. Allowing these emotional experiences to come forward requires a great deal of trust in the therapist. Thus, the therapist should ensure that her or she have a solid therapeutic alliance before attempting to facilitate experiencing too much. The therapist must also be comfortable with their own emotional experiencing. Because clients are often uncomfortable, the therapist should be familiar and comfortable with this inner terrain. Any discomfort on the therapist’s part may cause the client to withdraw or engage in negative self-judgments. It is unclear whether this comfort with experiencing comes with training or is more of an FIS type trait.
Thus, we see that therapists should use caution when facilitating experiencing. Possibly the trained therapists in the OUHRS proceeded to facilitate experiencing out of tradition, or because they assumed it would lead to good outcomes. Maybe training should focus on deciding when to facilitate experiencing. There is also the possibility that experiencing may be beneficial, yet the trained were not able to raise it enough and sacrificed some other therapy process in their efforts to raise experiencing. Something like this may have occurred, as the untrained therapists exhibited a trend towards having higher alliances than the trained therapists. This trend and its implications are discussed in the next section.

**Dangers of Technique Focus**

In this study, the trained therapists’ focus on experiencing may have hurt their ability to form alliances. A t-test comparison of client rated WAI scores at termination revealed a trend towards higher alliances for the untrained therapists, \( t(1, 42) = -1.881, p = .067 \). The trained therapists succeeded in facilitating experiencing, a process that was not related to outcome, but they were unable to develop higher alliances than the untrained therapists, and in fact that data was in the opposite direction, though not significantly. While it is unclear how these process variables interacted, one unsettling possible explanation should be explored. The therapists may have been striving for what has been accepted as common success strategy, experiencing, using techniques that may have hindered their ability to be successful.

There is some evidence from other research to support this interpretation. In the Vanderbilt II study the therapists received one year of manualized instruction in performing Time Limited Dynamic Psychotherapy (TLDP) (Henry et. al, 1993). The
training successfully changed technical interventions, although findings indicated that many therapists had not achieved a great deal of competence. Although technical behaviors changed in the expected direction, unexpected negative effects on some therapeutic behaviors were discovered. The therapists became less approving and supportive, less optimistic, and more authoritative and defensive. There may be some parallels between these findings and the current study.

In the OUHRS one group of therapists spent the previous two years learning specific therapy techniques through supervised practicum and classes. All of these students were newcomers to therapy, and thus were more likely focused on performing the correct technical interventions. The study took place under a principal investigator who identifies with the interpersonal theory and values humanistic approaches. Although the therapists were not instructed to perform any specific orientation, it is possible that they believed it was part of their job to facilitate experiencing. These therapists, who were evidently successfully trained to facilitate more experiencing than untrained therapists, may have focused on experiencing, yet lost something in process just as the Vanderbilt therapists who were following manuals did. The OUHRS trained therapists indeed showed a trend towards lower alliances than the untrained therapists and they were not able to outperform the untrained therapists on outcome measures. Possibly, training, at least in the beginning, can be a distraction to therapists. Beginning trainees may feel the presence of their supervisor sitting on their shoulder, evaluating their every move, judging the acuity of their interventions, and possibly inhibiting their ability to focus on the basics of the relationship.
Other studies have also shown that adherence to technique can sometimes have a negative impact on process variables and even outcome. In a study which aimed to differentiate variables unique to cognitive therapy and common factors such as the therapeutic alliance and emotional experiencing, the specific therapist technique of focusing on the impact of distorted cognitions was negatively correlated with outcome (Castonguay et al., 1996). Descriptive analyses of the sessions revealed that these therapists who applied the cognitive techniques failed to take the interpersonal context of the therapy into account. The researchers found that the therapists did not deal with alliance strains by searching for their possible source. Instead they returned to the cognitive model and its techniques and stated that the client’s hostility was due to his or her distorted thoughts. Thus, certain therapists in this study either did not know when to deviate from the model or lacked the interpersonal skills to develop an alliance which would allow the model to function.

*Ability to use and deliver common factors may not depend on training*

Several studies have found that nonprofessional therapists could not be distinguished from professional therapists on measures of “core conditions”. Carkhuff and Truax (1965) found that with limited training psychiatric aides were able to offer levels of warmth and empathy comparable to those offered by advanced graduate students and experienced therapists. Another study found that in initial “therapeutic” interviews, untrained college student volunteers were as warm, genuine, and empathetic as experienced psychiatrists and psychiatric residents (Pope, Nudler, VonKorff, & McGee, 1974).
FIS and Alliance left Standing

Did the High FIS, both trained and untrained provide something besides high experiencing that led to good outcomes. We know high FIS therapists had better alliances than the low FIS therapists, yet these alliances did not lead to significantly higher experiencing as has been the case in some studies. The alliance possibly effected outcomes through another pathway than experiencing in this study. This brings up interesting questions such as, “is the alliance curative in and of itself in short term therapy?” In fact the case studies revealed that some of the untrained therapists used non-traditional techniques such as frequent self-disclosure, edification, and lecturing. Yet they did always appear to build a strong alliance, and engage in curative processes, which may not be easily definable but were termed “buddy therapy” by a qualitative analysis team (Anderson et. al, 2001b).

A recent study also found a wide range in specific therapist behaviors, but a common thread of high alliances among high performing therapists. Luborsky (1997) studied 22 therapists’caseloads with drug-addicted and depressed clients and found important differences in the post-treatment outcomes among the therapists. Despite efforts to minimize differences between the therapists, the range of improvements for the clients of the 22 therapists varied from a slightly negative therapeutic impact to more than 80% improvement. Although there were client differences within therapists’ caseloads, it was the therapists’ capacities or skills that were responsible for change rather than differences in clients’ characteristics or diagnoses. Thus, better therapists did better with most of the clients they treated. The two best therapists in the study also used different techniques. One used mostly supportive techniques, whereas the other used mostly expressive
techniques. Even though the technical routes differed for these two highly effective therapists, they both showed one common trait, a strong working alliance with their clients. Thus, a strong alliance may allow therapists to take different routes to success. This is a further reason to not always make the specifics of therapy the object of research.

**Future studies**

Future studies could use a significant events model to code the sessions. Raters would evaluate sessions in order to find moments of likely higher experiencing. In this manner a greater range of experiencing scores may be possible. Other possible avenues to explore would be to conduct another study comparing trained and untrained individuals and ask them what they were trying to achieve. After each session clients and therapists would be interviewed to determine what they thought the most significant moment was. Therapists would be asked what they were trying to do. This would clarify if therapist training inculcated a desire to probe for depth. Or it could determine if both untrained and trained value experiencing, but the trained had the intervention skills to facilitate it better. If training still did not lead to improved outcome, one would again have to wonder about the way mental health professionals are trained.

Another possibility would be to film sessions and then emotion code them using a video coding system such as the one initially developed Gottman (1993) to code marital therapy and later refined by Giese-Davis (2002) to code the emotions of women living with metastatic breast cancer in group therapy. The systems can to take into account the facial movements (using Ekman’s Facial Affect Coding System), body language, voice tone and speech content. Experiencing is mostly a content based coding system and
interventions may bring about responses that result in higher experiencing codes without the client truly processing at greater depth.

A brief coding using VRM’s could be done to ascertain the different intervention pattern between trained and untrained therapists. It may be that trainees learn to talk like a therapist. This pattern seems different than a normal conversation (Stiles, 1979), and it most likely does have healing properties. However, part of this different style of conversation may be passed down through the generations of new therapists more out of tradition than empirical evidence.

Maybe the field needs to utilize a wider variety of process measures. The authors chose experiencing because it has a history and could be called reliable and valid. Maybe more qualitative exploratory research is necessary rather than simply trying to tease apart therapy using existing coding schemes and scales. Mahrer (1988) has stated that typical hypotheses testing helps us know less and less with greater and greater certainty. More studies such the one by Anderson et. al (2001), which helped flesh out some of these unique behaviors that the therapists with various combinations of FIS and training used, may be required

One possible option would be to take everything learned from this study and the original OUHRS and try to apply this knowledge to develop and study an effective therapy. The OUHRS revealed that the high FIS therapists developed higher alliances and that the untrained high FIS therapists practiced “buddy therapy”, which although unconventional, did seem to have some benefits (Anderson et. al, 2001b). We also have common factors models, which have been present for decades and posit that a concerned, helpful listener, with or without professional training, can lessen psychic distress if he or
she offers the client a warm, supportive relationship, and the client has confidence in the therapist’s abilities (Frank, 1973; Torrey, 1972; Truax & Mitchell, 1971).

Thus, a potential study would be to advise therapists to practice “Buddy Therapy”. This study could utilize high and low FIS groups and a brief manual describing this “Buddy” or “Alliance” therapy, in which the relationship or attending to the relationship would be the therapy. This would reveal how high and low FIS differ in their attempts and abilities to form alliances. It would also be interesting to examine what the therapists do in addition to “Buddy Therapy”. Many psychotherapy researchers would probably claim that therapy cannot be all common factors; some techniques or specifics must be used. However, if all the specifics are designed to foster the relationship, are they still in the realm of specifics or are they now common, once in the service of common factors? Is the distinction helpful in this case, and might the study help us gain a better understanding of exactly how the most basic of common factors, the alliance and FIS manifest themselves through the specifics of interventions, speech and behaviors?

Limitations

Some may say that the results of this study are limited because the sample of tape segments was not randomized. While, this would give greater generalizability to the results, it probably would not result in an overall increase in experiencing. Also many studies have shown that small segments are representative of the therapy hour as a whole (Kiesler et al., 1964; Mintz, 1969; Mintz & Luborsky, 1971; Suh et al., 1986; Weiss et al., 1988)

In order to determine if the 30-38th segment did represent the entire therapy hour, the first author assigned other time segments from therapy hours to the coders. He assigned
segments for minutes 14-22, 22-30, and 38-46 of various sessions. He did not assign all of these segments for the same session. In other words, he did not take one tape and assign the four segments to be coded. Rather he split the comparisons with the 30-38\textsuperscript{th} minute among a larger number of tapes, assigning one group of tapes for comparison with minutes 14-22, another for minutes 22-30, and another for minutes 38-46. Unfortunately this means that sessions 14-22, 22-30, and 38-46 cannot be easily compared. See table 11 for the correlations of the modes among the various segments and table 12 for the correlations among the peaks.

Table 11

Correlations between modes of different segments

<table>
<thead>
<tr>
<th>Segment</th>
<th>14-22\textsuperscript{nd} minute</th>
<th>22-30\textsuperscript{th} minute</th>
<th>38-46\textsuperscript{th} minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-38\textsuperscript{th} minute</td>
<td>.53*</td>
<td>.88**</td>
<td>.68*</td>
</tr>
<tr>
<td>n</td>
<td>22</td>
<td>19</td>
<td>10</td>
</tr>
</tbody>
</table>

* Correlation is significant at the .05 level (2-tailed).

** Correlation is significant at the .01 level (2-tailed).
Table 12

Correlations between peaks of different segments

<table>
<thead>
<tr>
<th>Segment</th>
<th>14-22nd minute</th>
<th>22-30th minute</th>
<th>38-46th minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-38th minute</td>
<td>.45*</td>
<td>.592**</td>
<td>.651*</td>
</tr>
<tr>
<td>n</td>
<td>22</td>
<td>19</td>
<td>10</td>
</tr>
</tbody>
</table>

* Correlation is significant at the .05 level (2-tailed).
** Correlation is significant at the .01 level (2-tailed).

Could coders have been biased and discerned who sounded trained and thus given these segments higher experiencing scores because it sounded more like appropriate therapy to them? Debriefing interviews of each coder suggested that this was not the case. Coders stated that they were not able to discern the outcomes of the study and were on the whole not able to tell who was trained and who was not.

Coders were debriefed to examine if there were any bias or problems in the coding process and to provide them with information about the study. The coders were asked a series of questions in an interview format with the principal investigator. Several questions were designed to estimate if any notions about hypotheses may have influenced their coding (See appendix for questions and coder responses). The coders seemed to focus on the task of coding the scale and not get overly involved in other aspects of the therapy such as who was doing good therapy. They seemed to benefit from consensus
coding and did not feel that having the principal investigator in the consensus meetings influenced their coding at all.

Overall, it appears that the coders were not able to guess the hypotheses. Several of the coders guessed hypotheses which were related to the study, but most of their guesses were about hypotheses which did not reach significance. Two coders hinted at hypotheses relating higher experiencing to higher outcomes. Coder A did guess that the hypotheses were related to the differences between trained and untrained. However, this coder also said that she was only able to guess who was trained about 10% of the time. Overall, looking at Table 13 it seems that all of the coders were in agreement that trained were facilitating higher experiencing than untrained.

Table 13
Mean experiencing scores for trained and untrained therapists

<table>
<thead>
<tr>
<th>Training Status</th>
<th>A</th>
<th>D</th>
<th>J</th>
<th>L</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mode</td>
<td>Peak</td>
<td>Mode</td>
<td>Peak</td>
<td>Mode</td>
</tr>
<tr>
<td>Trained</td>
<td>3.07</td>
<td>3.6</td>
<td>2.54</td>
<td>3.23</td>
<td>2.31</td>
</tr>
<tr>
<td>Untrained</td>
<td>2.0</td>
<td>3.0</td>
<td>2.0</td>
<td>3.0</td>
<td>1.91</td>
</tr>
</tbody>
</table>
The coders’ ability to decipher the training status of the therapists could have been empirically tested by having the coders rate several tapes. However, if this was done early in the study, it would have placed more emphasis on getting them to think about whether the therapist was trained or not. Also, if there were any experimenter effects in this study they probably would have been shown to favor the hypotheses about higher FIS therapists delivering higher experiencing. The first author has an affinity for some of the “buddy therapy” techniques used by the untrained therapists in the case analyses. He is more inclined to favor common factors and the FIS construct than he is to favor training programs. Therefore, he had no desire to see trained therapists rated as facilitating higher experiencing.

Conclusion

The above study delivered some interesting findings which were not fully expected. It attempted to elucidate how high FIS therapists influence the therapy process. It did not find an answer to this question, possibly indicating that the subtle ways that FIS manifest themselves need to be studied with more sophisticated tools, alliance scales excluded.

The study found evidence that training did influence experiencing, yet experiencing had no ties to outcome. Explanations were offered suggesting that experiencing may be technique related and therefore not such a common factor. Some professionals may take heart to learn that training is doing something, yet in this study the experiencing was not related to higher outcomes. Moreover, the trained therapists did not outperform the untrained in terms of the outcomes of their clients. That returns us to the original finding which showed that FIS contributed to higher alliances and outcomes. Maybe the trained therapists should reexamine why and how they are facilitating experiencing. Meanwhile
efforts will continue to validate the FIS construct and examine just how it contributes to therapeutic effectiveness.
REFERENCES


Bachelor, A. (1991). Comparison and relationship to outcome of diverse dimensions of
the helping alliance as seen by client and therapist. *Psychotherapy: Theory, Research and Practice, 28*, 534-549.


trials in psychotherapy research: Comment on Wampold et al. (1997).


pretreatment characteristics to the therapeutic alliance in diverse psychotherapies.


Hill, C. E., Thames, T. B., & Rardin, D. K. (1979). Comparison of Rogers, Perls, and


Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and


Therapist success and its determinants. *Archives of General Psychiatry, 42*, 602-611.

Luborsky, L., & Singer, B. (1975). Comparative studies of psychotherapies: Is it true that "Everyone has won and all must have prizes"? *Archives of General Psychiatry, 32*, 995-1008.


Pinsof, W. M. (1986). The process of family therapy: the development of the family
therapist coding system. In L. S. Greenberg & W. Pinsof (Eds.), The
Press.


client: Client process as a predictor of counselling skill acquisition. Canadian

relationship and its Impact: A study of psychotherapy with schizophrenics.
Madison: University of Wisconsin Press.

Rosenzweig, S. (1936). Some implicit common factors in diverse methods in
psychotherapy. American Journal of Orthopsychiatry, 6, 412-415.

rupture as a therapy event for empirical investigations. Psychotherapy: Theory,
Research and Practice, 27, 154-165.

Safran, J. D., Muran, J. C., & Wallner Samsta, I. (1994). Resolving therapeutic ruptures:
A task analytic investigation. In A. O. Horvath & L. S. Greenberg (Eds.), The
working alliance: Theory, research and practice (pp. 335-358). New York:
Wiley.

Schoeninger, D. W., Klein, M. H., & Mathieu, P. L. (1967). Sampling from recorded
therapy interview: patient experiencing ratings made with and without therapist


van der Veen, F. (1967). Basic elements in the process of psychotherapy: A research


APPENDIX A
Coder Debriefing Responses

Each coder’s answers is presented for each question. The first author attempted to capture their responses word for word. Coders responses is labeled by coder initials (A, D, J, L).

1. What do you think the hypotheses of this study were?
A: Different Levels Depending on Trained vs. Untrained. Clients might open up more.
L: The personality has more to do with skills (educational level) than being a friendly person. There is increased experiencing when the person is trained. Though not necessarily. I remember when you showed us that graph and it said that training doesn't necessarily matter that much.
J: I don't know. I didn't spend much time thinking about it. I just tried to do my job. If I had to guess, I would say that those who had higher experiencing would do better in therapy.
D: I didn't really think about it too much. Probably that the people who did better had higher experiencing.

2. Were you able to distinguish who was trained and untrained?
A: No, I had no idea who was trained.
L: Yes.
J: No, there was one instance when a therapist talked about being a chemistry student. But besides that, I couldn't tell.
D: Once in a while, they would mention something that would give it away, but I didn't really think about it.

3. What % of time could you guess?

A: I'd say 10%.

L: 70%

J: 5%

D: 10%

4. Were you able to focus on the coding task without considering any other therapy dynamics? (This question sometimes confused the coders. I would then clarify: Did you find yourself judging the therapist, the relationship or whether you thought it was good therapy, or could you simply follow the assigned task or rating the clients speaking turns?)

A: Yes, I applied the scale.

L: Yes, 100% of the time.

J: Yes.

D: Yes. I would just come in, do my tapes and leave. I didn't really engage in that much extra thought about the sessions. I had a lot of boring work this quarter, including data entry, and I got out of the habit of analyzing what I was doing.

5. Were you aware of which tapes were being double coded, and did this concern you?

A: No awareness or concern
L: No.

J: No

D: No, I had no idea.

6. Did you recognize stories within the dyads?

A: Sometimes recognized the story, but couldn't follow, 10% recognized, but not the session

L: I recognized one story but don't remember if it was in order or not.

J: Rarely. Once in a while things would seem familiar, but they seemed to be out of order and the clips were so short that we never hear the whole story.

D: There was this one therapist who I had to code twice and I hated her voice. Besides that I didn't really focus on stories.

7. Did consensus coding help you stay with the group and the manual?

A: Yes

L: In the beginning it was more helpful.

J: Yes, that was very helpful. I think we should have done more.

D: Yes, I found it to be very helpful, especially when we were training.

8. During consensus coding, was your opinion about coding swayed by the first author?

A: No, I don't think you had that much sway.

L: No, I think we kinda took over.

J: No, I don't think you tried to influence us. You seemed to mostly be asking us and the pointing us to the manual.
D: No, it seems like we had a good amount of control. I do remember that you would refer us to the manual about disagreements.